

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

DOUBLE TRANSMISSION UNIPARENTALE DE L'ADN MITOCHONDRIAL
CHEZ LES UNIONOIDAE : HÉRÉDITÉ, SÉLECTION ET ÉVOLUTION

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN BIOLOGIE
EXTENSIONNÉ DE L'UNIVERSITÉ DU QUÉBEC À MONTRÉAL

PAR
HÉLÈNE DOUCET-BEAUPRÉ

SEPTEMBRE 2012

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

À Philippe

« Je ne sais où va mon chemin.

Mais je marche mieux quand ta main serre la mienne. »

Alfred de Musset

*La recherche comporte et comportera toujours
une part importante d'activité créatrice.*

Pierre Joliot

REMERCIEMENTS

La concrétisation et l'accomplissement d'une telle thèse de doctorat ne peuvent se faire sans l'appui, les encouragements et l'ouverture de plusieurs personnes. Je tiens tout d'abord à remercier mon directeur de recherche, M. Pierre Blier (Université du Québec à Rimouski). Je te remercie sincèrement pour ta confiance, et pour le degré d'autonomie et de liberté que tu m'as laissée dans ce projet. Je salue tes idées audacieuses et ta créativité. Tu as su faire naître en moi une passion, ce qui n'est certes pas banal et ce qui est, à mon avis, le travail le plus important qu'un professeur puisse accomplir. Sans passion, il n'y a pas de persévérance et en ce sens je te dois beaucoup. J'ai pour toi un grand respect. J'ai toujours su que peu importe le déroulement du projet, je pouvais, en tout temps, te parler directement. J'apprécie la personne que tu es et les qualités humaines qui transcendent ton travail. Peut-être n'ai-je pas été des plus faciles à diriger, mais saches, qu'en tout temps, ton opinion comptait et compte toujours pour moi. J'espère que mon travail pourra te rendre fier et j'espère pouvoir continuer de collaborer avec toi. Après ces quelques années passées sous ta supervision, je constate que nous sommes scientifiquement beaucoup plus semblables que nous n'oserions nous l'avouer.

Je tiens à remercier tout particulièrement Dr Walter Randy Hoeh (University of Ohio), pour le vent de fraîcheur et de motivation que sa participation m'a amené dans ce projet. L'enthousiasme et l'ardeur que tu mets dans ton travail m'ont beaucoup inspirée. Tu es une des seules personnes que je peux entendre crier d'enthousiasme par courriel. Cette passion contagieuse a été pour moi d'un grand secours dans les moments un peu plus difficiles dans l'écriture de ma thèse. Je tiens également à te

remercier d'avoir mis le chapeau de codirecteur officieusement lorsque mon projet a changé de trajectoire vers quelque chose d'un peu plus génétique et génomique. Cette aide spontanée fut et continue d'être très appréciée. Je tiens également à te remercier de ta grande générosité et de ta confiance dans le partage de tes nombreuses données qui constituent le cœur de mon troisième chapitre. Un immense merci pour la pression, à peine déguisée, que tu as mise parfois pour faire avancer les choses un peu plus vite. Enfin, merci pour tes nombreuses et rigoureuses corrections, elles m'ont permis d'apprendre et de m'améliorer. Ta présence constante et virtuelle, puisque plus souvent par courriel, fut déterminante et fondamentale dans l'issue et la réussite de cette thèse. Un sincère merci.

Mes remerciements vont également au Dr Hans-Otto Pörtner (Alfred Wegener Institute, Germany) pour avoir accepté d'emblée de participer au projet initial. Merci pour le travail réalisé sur l'article « Thermal sensitivity of metabolic enzymes in subarctic and temperate freshwater mussels (*Bivalvia*: *Unionoida*) » publié dans *Journal of Thermal biology* en Janvier 2010 (cet article n'est pas présenté dans cette thèse). Merci d'avoir été une source d'inspiration majeure et un modèle de penseur, mon cœur sera toujours en physiologie.

Je tiens à exprimer mes remerciements aux membres du jury, qui ont accepté d'évaluer mon travail de thèse soit le Dr Jean Danyluk (président de jury, Université du Québec à Montréal), le Dr Nicolas Lartillot (membre externe, Université de Montréal) et le Dr Richard Cloutier (Professeur de l'UQAR, Université du Québec à Rimouski). Je suis très sensible à l'honneur que vous me faites en acceptant de participer à ce jury de thèse.

J'aimerais également remercier Dre France Dufresne pour m'avoir permis d'établir mon camp de base dans son laboratoire. Merci à Dre Jocelyne Pellerin de m'avoir initiée à la recherche. Le cheminement doctoral est le fruit de nombreuse discussion avec des amis et collègues de travail et de laboratoire. Je tiens à vous remercier tous et toutes du fond du cœur pour vos idées, vos encouragements, votre travail exemplaire et votre sympathie dans les moments plus creux. Vous avez alimenté la flamme et la passion pendant de nombreuses années. Parmi vous, un merci spécial à Sophie Breton, Simon Lamarre, Delphine Ditlecadet, Thomas Grunbaum, Richard Cloutier, Magella Guillemet, Hélène Lemieux, Josiane Cabana, Laurence Fisher-Rousseau, Roland Vergilino, Pierre Rioux, Nicolas Pichaud. Ceux et celles absents de cette liste ont tout de même une place de choix dans mon cœur et je vous remercie tout autant.

Les années de recherches cumulées dans cette thèse n'auraient pas été possibles sans le généreux soutien financier du Conseil de Recherche en Sciences Naturelles et en Génie du Canada (CRSNG/ NSERC), de même que le soutien financier du Fond Québécois de la Recherche sur la Nature et les Technologies (FQRNT).

Je tiens également à remercier toute ma famille pour son soutien constant tout au long de mes études et de mon doctorat. Merci d'avoir semé en moi cette graine d'intelligence et de curiosité essentielle à la poursuite d'étude graduée comme celle-ci. À mon frère, merci de ton indispensable participation à mon cinquième chapitre. J'ai toujours su qu'on finirait par travailler ensemble et je souhaite sincèrement que ce ne soit que le départ d'une belle aventure.

Parce que cette thèse est aussi une victoire personnelle sur la maladie, je remercie tous les médecins généralistes et spécialistes qui ont croisé mon chemin. Mon expérience des dernières années a réitéré ma foi en la recherche et en l'avancement des connaissances fondamentales. Je dois ma santé et ma survie à des personnes qui, comme moi, font des études avancées et qui ont à coeur de faire progresser la recherche. Je remercie les chercheurs de Schering-Plough. Puisque les dernières années furent également une longue réflexion sur la vie, la mort et sur ce qui a de l'importance, je dis aux chercheurs néophytes comme aux chercheurs aguerris, ne sacrifiez pas et ne négligez jamais ceux que vous aimez sur l'autel de votre réussite scientifique, cela n'en vaudrait pas la peine.

Enfin, mes plus grands remerciements vont bien évidemment à Philippe, mon amour, merci pour tout, pour ta présence, pour tes encouragements répétés, pour tes sacrifices, pour ta patience et ton soutien constant. Merci de m'avoir accompagnée scientifiquement et moralement. Je te dois tout.

TABLE DES MATIÈRES

LISTE DES FIGURES	xxiii
LISTE DES TABLEAUX.....	xxxi
LISTE DES ABRÉVIATIONS	xxxii
LISTE DES SYMBOLES.....	xxxvii
RÉSUMÉ	xxxviii
CHAPITRE I	
INTRODUCTION GÉNÉRALE.....	1
1.1 Les mitochondries.....	1
1.2 Génome mitochondrial.....	3
1.2.1 Organisation moléculaire	3
1.2.2 Réplication, transcription et traduction de l'ADN mitochondrial	8
1.2.2.1 Réplication	8
1.2.2.2 Transcription et traduction	10
1.2.3 Taux de mutation.....	10
1.2.4 Homoplasme et hétéroplasme	12
1.2.5 Héritabilité et transmission de l'ADNmt.....	14
1.2.6 Évolution de l'ADNmt.....	18
1.3 Double transmission uniparentale.....	21
1.3.1 L'exception qui confirme la règle	21
1.3.2 DUI : le modèle	22
1.3.3. Prévalence de la DUI chez les bivalves.....	23
1.3.4 Particularités de la DUI dans les différentes familles de bivalves	24
1.3.5 Évolution des génomes F et M	27

1.3.6 Objectif général du projet de recherche.....	29
1.3.7 Objectifs spécifiques du projet de recherche.....	30

CHAPITRE II

THE UNUSUAL SYSTEM OF DOUBLY UNIPARENTAL INHERITANCE OF MTDNA: ISN'T ONE ENOUGH?..... 33

Résumé	34
2.1 Abstract	35
2.2 Mitochondrial inheritance: rules for mussels.....	36
2.3 The DUI system unveiled.....	38
2.4 Molecular evolution of M and F mitochondrial genomes.....	45
2.5 When DUI breaks down: masculinization of F mtDNA.....	52
2.6 Origin and evolution of DUI.....	56
2.7 Potential adaptive evolution of M mtDNA	57
2.8 Concluding remarks.....	59
2.9 Acknowledgements.....	60
2.10 Glossary.....	60

CHAPITRE III

MITOCHONDRIAL PHYLOGENOMICS OF THE BIVALVIA (MOLLUSCA): SEARCHING FOR THE ORIGIN AND MITOGENOMIC CORRELATES OF DOUBLY UNIPARENTAL INHERITANCE OF MTDNA 63

Résumé	64
3.1 Abstract	65
3.2 Background.....	66
3.3 Results.....	72
3.3.1 Phylogenetic analysis	72
3.3.2 Genome structural features	75
3.3.3. Base composition and codon usage.....	81
3.3.4 Transfer RNA genes	81
3.3.5 Unassigned regions and putative control regions	82

3.3.6 Levels of intra- and interspecies sequence divergences	83
3.4 Discussion	85
3.4.1 Phylogenetic analysis	85
3.4.2 Genome structural features	86
3.4.3 tRNA Histidine	89
3.4.4 Extension of the M cytochrome c oxidase subunit II gene	89
3.4.5 Atp8 gene	90
3.4.6 NADH dehydrogenase subunits 4 and 4L genes	91
3.4.7 Base composition and codon usage	92
3.4.8 Transfer RNA genes	93
3.4.9 Unassigned regions and putative control regions	93
3.4.10 Levels of intra- and interspecies sequence divergences	94
3.5 Conclusions	96
3.6 Methods	97
3.6.1 DNA extraction, PCR amplification and sequencing	97
3.6.2 Gene annotation and analysis	100
3.7 Authors' contributions	104
3.8 Acknowledgements	105

CHAPITRE IV

<i>PYGANODON</i> (BIVALVIA: UNIONOIDA: UNIONIDAE) PHYLOGENETICS: A MALE- AND FEMALE-TRANSMITTED MTDNA GENOMIC PERSPECTIVE	106
Résumé	107
4.1 Abstract	108
4.2 Introduction	109
4.3 Materials and methods	113
4.3.1 Populations Sampled	113
4.3.2 DNA Extraction, amplification and sequencing	116

4.3.3 Phylogenetic reconstruction.....	117
4.4 Results.....	121
4.4.1 Phylogenetic analyses.....	132
4.4.2 Sequence-based species delimitation	136
4.4.3 Biogeography	137
4.4.4 Variation in F and M genome within and among clades	142
4.5 Discussion	144
4.5.1 Recent speciation of <i>P. cataracta</i> and <i>P. fragilis</i>	145
4.5.2 Non-monophyly/cryptic speciation in <i>P. grandis</i>	147
4.5.3 Sweep in <i>P. lacustris</i> M genome	149
4.6 Conclusion: evolution of M and F genomes in <i>Pyganodon</i>	150
4.7 Acknowledgements.....	152
CHAPITRE V	
POSITIVE SELECTION IN PATERNALLY INHERITED MITOCHONDRIAL GENOMES	153
Résumé	154
5.1 Abstract	155
5.2 Introduction.....	156
5.3 Materials and methods	160
5.3.1 Sequence alignment.....	160
5.3.2 Phylogeny reconstruction	162
5.3.3 Analysis of selective constraints	163
5.4 Results.....	166
5.5 Discussion	179
5.6 Conclusion.....	183
CHAPITRE VI	
CONCLUSION GÉNÉRALE	184

6.1 Principales propositions issues du chapitre II	185
6.2 Principaux résultats du chapitre III	186
6.3 Principaux résultats du chapitre IV	188
6.4 Principaux résultats du chapitre V	191
6.5 Généralisation et réflexions sur l'origine de la DUI	192
6.6 Contributions et limites du système de la DUI pour l'étude de la coévolution des génomes mitochondriaux et nucléaire	203
6.7 Signification des résultats pour la génétique mitochondriale et la biologie évolutive	205
6.8 Signification des résultats pour la biologie et l'écologie des Unionoïdés	209
6.9 Perspectives de recherches.....	213
APPENDICE A	
FIGURES SUPPLÉMENTAIRES DU CHAPITRE III.....	218
APPENDICE B	
TABLEAU SUPPLÉMENTAIRE DU CHAPITRE III.....	228
APPENDICE C	
FIGURES SUPPLÉMENTAIRES DU CHAPITRE V	232
BIBLIOGRAPHIE	237

LISTE DES FIGURES

Figure	Page
<p>1.1 Représentation schématique de la mitochondrie et du métabolisme cellulaire des substrats tels que le glucose et les acides gras (flèche verte sur la figure) dans la chaîne de transport des électrons. Figure inspirée de Krauss, Zhang et Lowell (2005).</p>	3
<p>1.2 (a) ADN mitochondrial humain type codant pour 13 polypeptides entrant dans la composition de (b) la chaîne respiratoire mitochondriale et l'ATP synthase. Tirée de Blier, Dufresne et Burton (2001).</p>	7
<p>2.1 DUI phylogeny. The mitochondrial genome evolutionary relationships for the DUI-containing taxa showing that the male (M) and female (F) genomes in the freshwater mussels, the Unionidae, have been stable whereas there is evidence for occasional masculinization events (i.e. F genomes giving rise to M genomes) in the families Veneridae and Mytilidae. Topology and branch lengths are based on analysis of 199 amino acids from <i>cox1</i> and posterior probabilities ($\times 100$) > 90 are displayed. The taxonomic and sex-specific transmission affiliations of the individual sequences are indicated at the right. The Mr. Bayes analysis (Huelsenbeck and Ronquist, 2001) used the Jones, Taylor and Thornton (JTT) model of amino acid substitution (Jones, Taylor and Thornton, 1992) and was run for 10 million generations with eight chains. Overall, 10 000 trees were saved during the course of the analysis but only the last 9000 were used to construct the consensus tree (i.e. the first 1000 trees were discarded as burn-in). The original nucleotide sequences (species and accessions numbers) were translated with the <i>Drosophila</i> mitochondrial genetic code and used according to Refs (Beagley, Taylor and Wolstenholme, 1997; Hoeh <i>et al.</i>, 1998; Hoeh, Stewart and Guttman, 2002; Boore, Medina and Rosenberg, 2004; Mock <i>et al.</i>, 2004; Riginos <i>et al.</i>, 2004; Mizi <i>et al.</i>, 2005; Breton <i>et al.</i>, 2006).....</p>	39
<p>Box 1 Figure I Proposed genetic control of mitochondrial transmission under the DUI system. (a) Male and female gametes contribute to the mitochondrial population of the fertilized egg (b) Mother-dependant genetic models (Saavedra, Reyero and Zouros, 1997; Zouros, 2000; Dalziel and Stewart, 2002; Cogswell, Kenchington</p>	

and Zouros, 2006) predict that mothers carrying the Z allele will produce eggs with the Z factor, allowing the retention of sperm mitochondria and their subsequent aggregation in a single blastomere in embryos destined to become males. By contrast, zz mothers will produce eggs without the Z factor in which sperm mitochondria will be dispersed and/or lost, and embryos will become females.

45

Box 2. Figure 1 Divergences in model systems for the study of co-evolution. Amino acid sequence differences of mtDNA-encoded peptides in model systems for analyses of cytonuclear co-evolution in which disruption of respiratory chain function has been observed. Human cells with mtDNA from chimpanzee and gorilla had impaired oxygen consumption (i.e. decreased, on average, by 20%, 34%, and 27%, respectively, compared with the human cells and/or human mtDNA control) that was attributable to a marked deficiency in respiratory chain complex I (Kenyon and Moraes, 1997; Barrientos, Kenyon and Moraes, 1998). Activities of complexes I, III, and IV in the *Rattus* xenocybrid were 44%, 37%, and 78% of control mean, respectively, whereas the activity of complex III was 2% of control cybrid in the *Otomys* xenocybrid (McKenzie and Trounce, 2000; McKenzie *et al.*, 2003). COX activities were decreased as mitochondria from a population were moved to a purer and purer nuclear background from a distant population (Burton, Rawson and Edmands, 1999; Ellison and Burton, 2006).*, Data not available.

51

- 3.1 Bayesian inference majority-rule tree of bivalve mt genome relationships based on an analysis using the M3 codon substitution model and a nucleotide alignment of 12 mitochondrial protein-coding genes (atp8 excluded). Numbers above an internal branch, from top to bottom, indicate nodal support values from BI, ML and MP nucleotide-based analyses, respectively. Numbers below an internal branch, from top to bottom, indicate nodal support values from BI, ML and MP amino acid-based analyses, respectively. Only nodal support values > 50% are presented. An asterisk above an internal branch indicates that all three nucleotide-based nodal support values are 100; an asterisk below an internal branch indicates that all three amino acid-based nodal support values are 100. Branch lengths reflect substitutions per site and the taxonomic and gender-specific transmission affiliations of the individual sequences are indicated at the right. All phylogenetic analyses

- strongly indicate that the unionoids represent the basal lineage for the bivalve taxa represented in this analysis. 73
- 3.2 ML-based (A) and Dollo parsimony-based (B) ancestral character state reconstructions of presence/absence of DUI on a species-level tree derived from the tree presented in Figure 3.1. The ML-based reconstruction indicates three separate origins of DUI while the Dollo parsimony-based reconstruction indicates a single origin of DUI followed by three reversals to SMI..... 74
- 3.3 Gene maps of the M and F mitochondrial genomes of *Venustaconcha ellipsiformis*, *Pyganodon grandis*, *Inversidens japonensis* and *Quadrula quadrula*. Protein and rRNA genes are named as in the text while tRNA genes are abbreviated by the one-letter code of the corresponding amino acid [(L1 =trnL (cua), L2 = trnL (uaa), S1 = trnS (aga), and S2 = trnS (uaa)]. Genes positioned inside the plain line are encoded on the heavy strand and genes outside the line are encoded on the light strand. Atp8* (= genomes lacking full size atp8 gene). Black arrows on the *V. ellipsiformis* M genome indicate regions that differ between male- and female-transmitted genomes and the arrow on the *I. japonensis* F genome indicates the region with a gene order distinct from that of the other figured F genomes. The circular gene maps of the genomes were drawn by GenomeVx (Conant and Wolfe, 2008) followed by manual modification. 76
- 4.1 Sampling sites of *Pyganodon* spp. (black ▼) and geographical occurrence of *Pyganodon* spp. (yellow ■) based on Global biodiversity information facility (GBIF; Biodiversity occurrence data provided by: Canadian Museum of Nature, EMAN Provider, New Brunswick Museum, National Museum of Natural History and University of Kansas Biodiversity Research Center (Accessed through GBIF Data Portal, data.gbif.org, 2010-03-26). Sites are numbered as in Table 4.1. The maximum extent of the proglacial lakes (dashed line) is outlined. 115
- 4.2 Bayesian tree of the F and M *Pyganodon* complex inferred from cox1 and cox2. ML bootstrap values (below) and BI posterior probabilities (x100, above) are indicated at the branches. Boxes indicate to which species the specimens were identified morphologically prior to the genetic analysis. 122

- 4.3 Bayesian majority-rule consensus tree of the F *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represent the morphospecies, the second represent the genome (F vs. M). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. In the last column, two tones of gray represent a connected network divided in two distinct subnetworks. 124
- 4.4 Bayesian majority-rule consensus tree of the F *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represent the morphospecies, the second represent the genome (F vs. M) and the third represents a subclade (A or B). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. 126
- 4.5 Bayesian majority-rule consensus tree of the F *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M) and the third represent a subclade (A, B or C). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. 128
- 4.6 Bayesian majority-rule consensus tree of the M *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column

- refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. In the last column, two tones of gray represent a connected network divided in two distinct subnetworks. 130
- 4.7 Bayesian majority-rule consensus tree of the M *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks..... 132
- 4.8 Bayesian majority-rule consensus tree of the M *Pyganodon* complex inferred from *cox1* and *cox2*. ML bootstrap values (below) and BI posterior probabilities (x100, above) are indicated at the branches. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M) and the third (if necessary) represents a subclade (A, B or C). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. In the last column, two tones of gray represent a connected network divided in two distinct subnetworks. 134
- 4.9 Geographic distribution of *Pyganodon cataracta* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). Color codes and clade names follow Figures 4.3 and 4.6. An empty half circle indicates that no data exist for that location. Sites are numbered as in Table 4.1. Symbol (C) Data from Cyr *et al.* (2007). The purple zone indicates area of the 36 localities from Cyr *et al.* (2007). 139

- 4.10 Geographic distribution of *Pyganodon fragilis* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). An empty half circle indicates that no data exist for that location. Color codes and clade names follow Figures 4.3 and 4.6. Sites are numbered as in Table 4.1. Symbol (C) means data from Cyr *et al.* (2007)..... 140
- 4.11 Geographic distribution of *Pyganodon lacustris* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). An empty half circle indicates that no data exist for that location. Color codes and clade names follow Figures 4.4 and 4.7. Sites are numbered as in Table 3.1. Symbol (C) means data from Cyr *et al.* (2007)..... 141
- 4.12 Geographic distribution of *Pyganodon grandis* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). An empty half circle indicates that no data exist for that location. Color codes and clade names follow Figures 4.5 and 4.8. Sites are numbered as in Table 3.1. Symbol (C) means data from Cyr *et al.* (2007)..... 142
- 5.1 DUI phylogeny. The mitochondrial genome evolutionary relationships for the DUI-containing taxa showing that the male (M) and female (F) genomes in the freshwater mussels, the Unionidae, have been stable whereas there is evidence for occasional masculinization events (i.e. F genomes giving rise to M genomes) in the families Veneridae and Mytilidae. Topology and branch lengths are based on analysis of 199 amino acids from *cox1* and posterior probabilities ($\times 100$) > 90 are displayed. The taxonomic and sex-specific transmission affiliations of the individual sequences are indicated at the right. The Mr. Bayes analysis (Huelsenbeck and Ronquist, 2001) used the Jones, Taylor and Thornton (JTT) model of amino acid substitution (Jones, Taylor and Thornton, 1992) and was run for 10 million generations with eight chains. Overall, 10 000 trees were saved during the course of the analysis but only the last 9000 were used to construct the consensus tree (i.e. the first 1000 trees were discarded as burn-in). The original nucleotide sequences (species and accessions numbers) were translated with the *Drosophila* mitochondrial genetic code and used according to Refs (Beagley, Taylor and Wolstenholme, 1997; Hoeh *et al.*, 1998; Hoeh, Stewart and Guttman, 2002; Boore, Medina and Rosenberg,

2004; Mock <i>et al.</i> , 2004; Riginos <i>et al.</i> , 2004; Mizi <i>et al.</i> , 2005; Breton <i>et al.</i> , 2006; Doucet-Beaupré <i>et al.</i> , 2010).	167
5.2 Proportion of sites in ω classes (negatively selected, neutral or positively selected) according to their branch classes (background (back) vs. foreground (fore)) and null vs. alternative (alt) hypothesis in the unionoid genome (A: Male transmitted genome; B: Female transmitted genome). LRT: likelihood ratio test; p : χ^2 probability; α_{BC} : Bonferroni corrected; ω_2 : maximum likelihood point estimate for ω_2 on the foreground branch for the fraction of sites under positive selection.	170
5.3 Proportion of sites in ω classes (negatively selected, neutral or positively selected) according to their branch classes (background (back) vs. foreground (fore)) and null vs. alternative (alt) hypothesis in the veneroid genome (A: Male transmitted genome; B: Female transmitted genome). LRT: likelihood ratio test; p : χ^2 probability; α_{BC} : Bonferroni corrected; ω_2 : maximum likelihood point estimate for ω_2 on the foreground branch for the fraction of sites under positive selection.	171
5.4 Proportion of sites in ω classes (negatively selected, neutral or positively selected) according to their branch classes (background (back) vs. foreground (fore)) and null vs. alternative (alt) hypothesis in the mytiloid genome (A: Male transmitted genome; B: Female transmitted genome). LRT: likelihood ratio test; p : χ^2 probability; α_{BC} : Bonferroni corrected; ω_2 : maximum likelihood point estimate for ω_2 on the foreground branch for the fraction of sites under positive selection.	172
5.5 Posteriori probability mass functions (PMF) of π marginalized over foreground branch sites and whole background for targeted lineages where likelihood ratio test gives significant results of positive selection. Dots show the discretization $\omega_0 \sim U(0,1)$, $\omega_1 \sim \delta(1)$, $\omega_2 \sim U(1,31)$ and steps are guide to the eye. The rightmost figure show the same PMFs on a logarithmic scale to emphasize the small but non-zero probability mass of sites to be under positive selection.	175
5.6 Bayesian empirical Bayes (BEB) probability for each site to be either negatively selected ($\Pr(\omega_0)$; red), neutral ($\Pr(\omega_1)$; blue) or positively selected ($\Pr(\omega_2)$; green) in targeted lineages where likelihood ratio test give significant results of positive selection.....	176

- 5.7 Bayesian empirical Bayes (BEB) probability for each site to be positively selected ($\Pr(\omega_2) \geq 0.95$) in targeted lineages where likelihood ratio test give significant results of positive selection..... 177
- 6.1 Les destinées évolutives possibles des gènes dupliqués inclus la néofonctionnalisation, la subfonctionnalisation, la redondance et la non-fonctionnalisation. Ces voies d'évolution sont la conséquence de la sélection naturelle, des mutations et de processus épigénétiques qui agissent sur les régions de régulation et sur les régions codantes des gènes dupliqués. Dans cette figure, une protéine est représentée par une paire de pince, chacune des branches de la paire représente une sous-fonction distincte de l'outil. La néofonctionnalisation transforme une des copies en un nouvel outil (un marteau) avec une nouvelle fonction. La subfonctionnalisation dégrade une des branches complémentaires de la protéine dupliquée (grise pâle) rendant l'activité combinée des deux paralogues nécessaire pour accomplir la fonction ancestrale. La non-fonctionnalisation dégrade complètement une des protéines dupliquées. Inspirée de Evans (2008). 196
- 6.2 Diagramme conceptuel de l'origine de la DUI et de son évolution selon les modèles de duplication génique et génomique connus. 202

LISTE DES TABLEAUX

Tableau	Page
1.1 Comparaison entre les génomes nucléaires et mitochondriaux humains. Tableau modifié de (Taylor and Turnbull, 2005).	4
2.1 Complete mitochondrial genomes of species with DUI.	47
3.1 Species and GenBank accession numbers of the sequences used in this study for phylogenetic analyses.....	70
3.2 Main structural features of the female- and male-transmitted mitochondrial genomes of <i>Venustaconcha ellipsiformis</i> , <i>Pyganodon grandis</i> , <i>Inversidens japonensis</i> and <i>Quadrula quadrula</i>	78
3.3 Unassigned regions of the female- and male-transmitted mitochondrial genomes of <i>Venustaconcha ellipsiformis</i> , <i>Pyganodon grandis</i> , <i>Inversidens japonensis</i> and <i>Quadrula quadrula</i>	83
3.4 Intra- and interspecies comparisons for 12 mitochondrial protein-coding genes: average pairwise sequence divergence.....	84
4.1 Description and geographical coordinates of sampling sites for <i>Pyganodon lacustris</i> , <i>P. cataracta</i> , <i>P. fragilis</i> and <i>P. grandis</i> and numbers of individual per morphospecies sampled (see Figure 4.1 and Figures 4.9 to 4.12) documented in each site and morphospecies.....	113
4.2 Morphospecies and divergence in percentage (%) between F.....	134
4.3 Species divergence in percentage (%) between M.....	135
4.4 Hierarchical AMOVA for the M and F data sets from <i>Pyganodon</i> morphospecies/clades.....	143
5.1 Names and GenBank accession numbers of the 34 species used in our evolutionary analysis of bivalve mitochondrial genomes.	161
5.2 Summary of branch-site test for positive selection along the 6 <i>a priori</i> selected lineages.....	168

LISTE DES ABRÉVIATIONS

A	adenine / adénine
a.a./aa	amino acid / acide aminé
ADN / DNA	Acide désoxyribonucléique / Deoxyribonucleic acid
ADNmt / mtDNA	ADN mitochondrial / Mitochondrial DNA
ADNnu / nuDNA	ADN nucléaire / Nuclear DNA
ARN / RNA	Acide ribonucléique / Ribonucleic acid
ARNr / rRNA	ARN ribosomal / ribosomal RNA
ARNt / tRNA	ARN de transfert / transfer RNA
ATP	Adénosine 5'triphosphate / Adenosine 5'triphosphate
BEB	Bayes empirical Bayes method / méthode Bayes empirique Bayes
BER	base excision repair / réparation par excision de base
BI	Bayesian Inference / Inférence Bayésienne
bp	base pair / paire de base
C	cytosine / cytosine
Coenzyme Q	Ubiquinone / ubiquinone
COX	Cytochrome <i>c</i> oxidase
CS	Citrate synthase
CYC	Cytochrome <i>c</i>
d.f.	degree of freedom / degré de liberté

D-loop	displacement loop / boucle de séparation
d_N	numbers of non-synonymous substitutions / nombre de substitutions synonymes
dNTP	Deoxyribonucleotide triphosphate / Désoxyribonucléotide triphosphate
d_S	numbers of synonymous substitutions / nombre de substitutions synonymes
DUI	Doubly uniparental inheritance / double transmission uniparentale
EDTA	Ethylenediaminetetraacetic acid / acide éthylène-diamine-tétraacétique
ETS	Electron transport system / système de transport des électrons
$FADH_2$	Reduced flavin adenine dinucleotide / Flavine adénine dinucléotide réduite
G	guanine / guanine
GMYC	general mixed Yule-coalescent method / méthode générale Yule-coalescence
GTR	Generalised time reversible model of amino acid substitution / modèle de substitution d'acides aminés réversible généralisé.
JTT	Jones, Taylor and Thornton model of amino acid substitution / modèle de substitution d'acides aminés de Jones, Taylor et Thornton.
K_a	rate of non-synonymous substitution / taux de substitutions non-synonymes
kb	kilobase

K _s	rate of synonymous substitution / taux de substitutions synonymes
LRT	Likelihood ratio test / test du rapport de vraisemblance
MgCl ₂	Magnesium chloride / chlorure de magnésium
ML	Maximum likelihood / maximum de vraisemblance
MP	Maximum parsimony / maximum de parsimonie
NADH/ND	Nicotiamide adenine dinucleotide / Nicotiamide adénine dinucléotide
Ne	Effective population size / taille de population effective
NRF	Nuclear respiratory factor / facteur respiratoire nucléaire
OXPHOS	Oxidative phosphorylation / phosphorylation oxidative
PCR	Polymerase chain reaction / réaction de polymérisation en chaîne
Polg	DNA polymerase subunit gamma / sous-unité gamma de l'ADN polymérase
SMI maternelle	Strict maternal inheritance / transmission strictement
sp./spp.	Species / espèce(s)
SSB	Single-strand binding protein / protéine fixant l'ADN simple brin
T	thymine / thymine
TFAM	Transcription factor A, mitochondrial / Facteur de transcription A mitochondrial
trnA	tRNA for alanine / ARNt pour l'alanine
trnC	tRNA for cysteine / ARNt pour la cystéine

trnD	tRNA for aspartate / ARNt pour l'aspartate
trnE	tRNA for glutamate / ARNt pour le glutamate
trnF	tRNA for phenylalanine / ARNt pour la phénylalanine
trnG	tRNA for glycine / ARNt pour la glycine
trnH	tRNA for histidine / ARNt pour la histidine
trnI	tRNA for isoleucine / ARNt pour l'isoleucine
trnK	tRNA for lysine / ARNt pour la lysine
trnL	tRNA for leucine / ARNt pour la leucine
trnM	tRNA for methionine / ARNt pour la méthionine
trnN	tRNA for asparagine / ARNt pour l'asparagine
trnP	tRNA for proline / ARNt pour la proline
trnQ	tRNA for glycine / ARNt pour la glycine
trnR	tRNA for arginine / ARNt pour l'arginine
trnS	tRNA for serine / ARNt pour la sérine
trnT	tRNA for threonine / ARNt pour la thréonine
trnV	tRNA for valine / ARNt pour la valine
trnW	tRNA for tryptophane / ARNt pour le tryptophane
trnY	tRNA for tyrosine / ARNt pour la tyrosine
U	Uracil / uracile
WAG	Whelan and Goldman model of amino acid substitution / modèle de substitution d'acides aminés de Whelan and Goldman.

LISTE DES SYMBOLES

α	alpha / alpha
γ	gamma
δ / Δ	delta / delta
κ	kappa
μM	micromolar / micro molaire
pg	picogram / picogramme
π	pi
χ^2	chi squarre / chi carré
ω	omega

RÉSUMÉ

Les mitochondries, organites essentiels à la production d'énergie, possèdent leur propre matériel génétique : l'ADN mitochondrial (ADNmt). Chaque produit des gènes de l'ADNmt interagit avec des protéines codées par l'ADN nucléaire pour assurer la respiration mitochondriale. Dans le règne animal, l'ADNmt est transmis exclusivement de façon maternelle. La transmission maternelle permettrait d'éviter les conflits intergénomiques. Il existe un seul système mitochondrial qui transgresse singulièrement les lois de la transmission et de la génétique de l'ADNmt, il s'agit de la double transmission uniparentale [doubly uniparental inheritance (DUI)]. L'existence de la DUI a été démontrée chez sept familles de mollusques bivalves, dont les moules marines : Mytilidae; les palourdes marines : Veneridae, Solenidae et Donacidae et les moules d'eau douce: Unionidae, Margaritiferidae et Hyriidae. Les espèces possédant ce système atypique sont caractérisées par la présence de deux ADNmt distincts qui sont hérités de façon maternelle (ADNmt F) et paternelle (ADNmt M). Typiquement, les femelles sont homoplasmiques et contiennent uniquement l'ADNmt F tandis que les mâles sont hétéroplasmiques. Chez ces derniers, l'ADNmt F domine dans les tissus somatiques alors que la gonade contient presque exclusivement l'ADNmt M. Les divergences observées entre les ADNmt F et M chez les moules peuvent atteindre ~50 % en acides aminés. Ce haut niveau de divergence observée entre les génomes M et F, coexistant à l'intérieur d'un même individu soulève de nombreuses questions d'ordre évolutives. La double transmission uniparentale représente, certes, un système atypique, mais les descriptions, les comparaisons et les analyses qu'elle permet contribuent à développer une image plus exacte de l'évolution non seulement au niveau de l'évolution de la DUI, mais aussi au niveau de l'évolution du génome mitochondrial dans son ensemble.

L'objectif de la présente thèse est d'utiliser une approche comparative de données moléculaires pour mettre en évidence l'évolution des génomes mitochondriaux mâle et femelle du système de la double transmission uniparentale chez les moules unionoïdes. Chez ce groupe, la double transmission uniparentale semble évolutivement plus stable (aucun événement de recombinaison ou de masculinisation n'a été recensé) et la divergence entre les ADNmt M et F est significativement plus élevée que chez les taxons marins. Le chapitre II est une revue de littérature sur le système de la double transmission uniparentale. Les observations qui ont menées à la découverte de la DUI y sont présentées de même que le modèle et les mécanismes de la DUI. Cette revue propose la DUI comme un système

modèle qui pourrait permettre d'acquérir de nouvelles connaissances sur les interactions et la coadaptation des génomes nucléaires et mitochondriaux.

L'objectif du chapitre III était de séquencer, d'annoter et de publier pour la toute première fois des génomes mâles complets d'Unionoïdés. Six nouveaux génomes mitochondriaux complets d'espèces d'Unionoïdés, soit le génome F et M de *Venustaconcha ellipsiformis* (Unionoida: Unionidae: Ambleminae: Lampsilini), *Pyganodon grandis* (Unionoida: Unionidae: Unioninae: Anodontini) et *Quadrula quadrula* (Unionoida: Unionidae: Ambleminae: Quadrulini) ont été séquencés avec succès et les caractéristiques génomiques propres à chaque lignée mâle ont été présentées et ont fourni un contexte pour les comparaisons des génomes mitochondriaux chez les lignées de bivalves possédant- et ne possédant pas la DUI. La position basale des Unionoïdés à l'intérieur des autolamellibranches ainsi que la probable origine unique de la DUI suggère que la DUI serait apparue chez un ancêtre des autolamellibranches et aurait été subséquentement perdue chez plusieurs descendants. Les caractéristiques de la DUI observées chez les Unionoïdés s'approcheraient donc davantage de l'état ancestral de la DUI.

Le chapitre IV est une étude phylogéographique du genre *Pyganodon* (Unionidae, Bivalvia) dans le nord-est de l'Amérique du Nord. Le genre *Pyganodon* a été choisi entre autres en raison de sa richesse taxonomique et de sa vaste répartition géographique dans le nord-est de l'Amérique du Nord. Cette analyse phylogéographique est basée sur l'étude du polymorphisme de gènes des ADNmt M et F ce qui a permis de comparer et confronter les signaux phylogénétiques rendus par les génomes M et F à l'intérieur de ce groupe. L'exploration de la distribution de la variation génétique entre et à l'intérieur des populations et des espèces de *Pyganodon* a permis de mettre en évidence la complexité du système, mais aussi de distinguer les processus démographiques des processus sélectifs. Ce chapitre renforce l'hypothèse de l'action de la sélection adaptative sur le génome mitochondrial mâle.

Très tôt après la découverte du système de la DUI, un possible rôle de la sélection positive comme pression sélective majeure a été soupçonnée. Malgré l'accumulation de certains indices, un tel processus n'avait jamais été démontré. Dans le chapitre V de la présente thèse, en utilisant les données recueillies et cumulées dans le chapitre III soit l'ensemble des gènes mitochondriaux codants (sauf l'ATP8) de 29 espèces de mollusque bivalves de différentes familles, un test de sélection positive en maximum de vraisemblance a été effectué. Les résultats de ce test suggèrent que de nombreux sites d'acides aminés sont positivement sélectionnés sur le

génomique mâle des Unionoïdés. Quelques sites seraient également positivement sélectionnés dans l'ADNmt F des Unionoïdés et témoignent probablement de la coévolution cytonucléaire. Le test n'a pas mis en évidence de sélection positive dans les ADNmt M ni F de Mytilidés et confirment que les forces évolutives n'agissent pas de manière identique dans les différents groupes de bivalves possédant la DUI. Les résultats du test révèlent que des sites d'acides aminés seraient également positivement sélectionnés dans le troisième groupe de bivalves testé soit les Vénéridés. Toutefois, la concentration de ces sites sur certains gènes spécifiques de l'ADNmt F pourrait témoigner de l'action de la sélection positive suite à une réorganisation structurale du génome. Compte tenu du profond impact que peut avoir l'ensemble des résultats de ce chapitre sur notre compréhension actuelle des forces sélectives agissant sur le génome mitochondrial, la prudence est toutefois de mise dans leurs interprétations.

Mots clés : ADN mitochondrial, génomique comparative, phylogéographie, sélection positive, double transmission uniparentale, bivalves

CHAPITRE I

INTRODUCTION GÉNÉRALE

1.1 Les mitochondries

Les mitochondries (du grec mitos, filament et chondros, granule) sont des organites cellulaires ovalaires délimités par une double membrane : une membrane interne entoure la matrice mitochondriale et une membrane externe délimite l'espace intermembranaire (Figure 1.1). La morphologie de la mitochondrie change selon le contexte cellulaire. Les mitochondries peuvent se retrouver sous la forme de petites unités individuelles fragmentées ou fusionnées pour former un large réseau de mitochondries allongées et interconnectées (Chan, 2006). Cette organisation structurale permet à la mitochondrie d'assurer différentes fonctions métaboliques. Les mitochondries produisent 90 à 95% de l'énergie utilisable (Magnus et Keizer, 1998) par les cellules animales grâce à un processus appelé phosphorylation oxydative (Hatefi, 1985). Cette énergie, produite sous forme d'adénosine triphosphate (ATP), résulte d'un transport d'électrons le long de la chaîne respiratoire située sur la membrane interne de la mitochondrie, avec une consommation d'oxygène (Saraste, 1999) (Figure 1.1).

La fonction principale des mitochondries est d'assurer l'essentiel de la production énergétique de la cellule, mais elles interviennent également dans de nombreux processus biologiques fondamentaux tels que la cétogenèse, l'uréogénèse, l'homéostasie calcique, l'homéostasie du fer, la production d'espèces réactives de l'oxygène, la synthèse des stéroïdes, la thermogénèse et l'apoptose c.-à-d., mort cellulaire programmée (Green et Reed, 1998; Skulachev, 1999; Martinou, Desagher et Antonsson, 2000; Parekh, 2003).

Selon la théorie endosymbiotique en série (Sagan, 1967; Margulis, 1981), la mitochondrie naquit de l'endosymbiose d'une α -protéobactérie dans les cellules eucaryotes primitives. Certains auteurs soutiennent que l'origine des eucaryotes et l'origine des mitochondries constitueraient un seul et même événement (Lane et Martin, 2010). Toutefois, nos connaissances sur comment, au début de l'évolution des eucaryotes, une bactérie autonome libre est devenue un organe semi-autonome demeurent limitées et controversées. La transition entre ces deux états a probablement impliqué quatre étapes : (i) la cellule-hôte primitive a capturé et englouti la bactérie par endocytose (phagocytose); (ii) la cellule-hôte contrôla la division cellulaire et la multiplication de la bactérie endosymbiotique (le préorgane) avec sa machinerie de division; (iii) une partie des gènes bactériens fut transférée au génome nucléaire de la cellule-hôte; (iv) avec le développement de la reproduction sexuée, le génome de la cellule-hôte régule l'évolution du génome de l'organe via l'héritabilité uniparentale (Kuroiwa, 2010a, 2010b). En situation de reproduction sexuée, la transmission uniparentale du génome mitochondrial semblerait être une méthode efficace pour éviter l'évolution de l'organe à travers la recombinaison des gènes mitochondriaux des deux parents (Kuroiwa, 2010a). Bien qu'au cours de l'évolution la plupart des gènes procaryotes furent transférés au génome nucléaire (Timmis *et al.*, 2004), les mitochondries ont conservé quelques vestiges de leurs ancêtres endosymbiontes, le plus notable sur le plan évolutif étant certainement qu'elles possèdent leur propre matériel génétique, l'ADN mitochondrial (ADNmt).

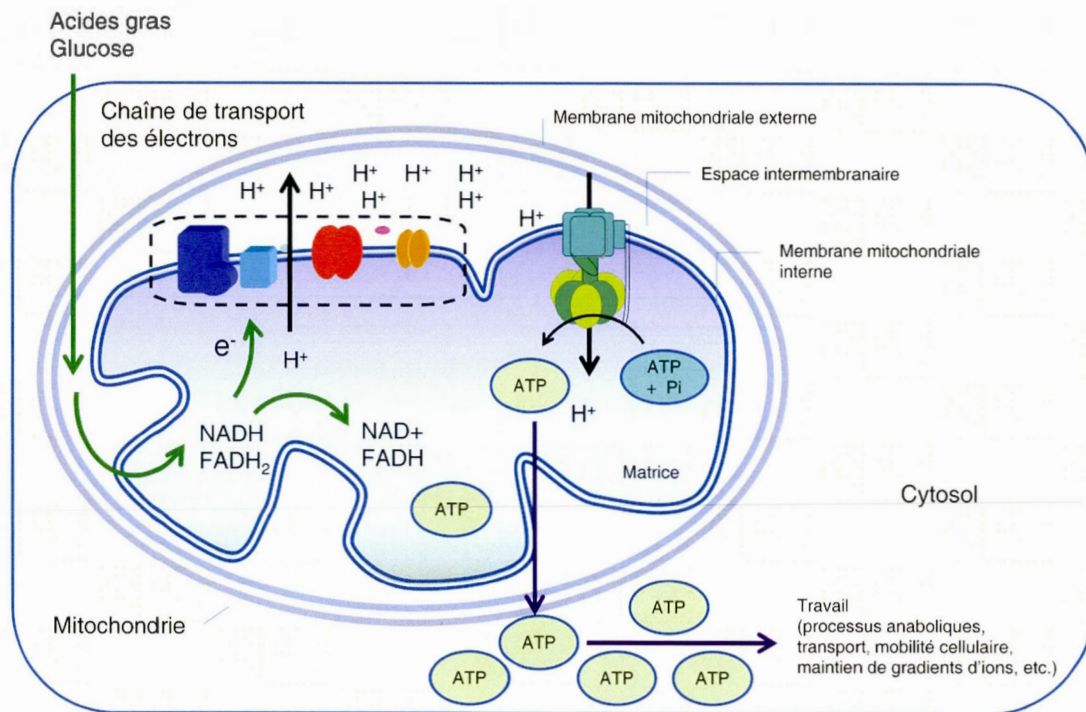


Figure 1.1. Représentation schématique de la mitochondrie et du métabolisme cellulaire des substrats tels que le glucose et les acides gras (flèche verte sur la figure) dans la chaîne de transport des électrons. Figure inspirée de Krauss, Zhang et Lowell (2005).

1.2 Génome mitochondrial

1.2.1 Organisation moléculaire

Le génome mitochondrial montre une grande variabilité en termes de structure, de contenu, d'organisation, et de mode d'expression des gènes dans les divers groupes d'organismes. Néanmoins, plusieurs caractéristiques sont communes à la majorité des génomes mitochondriaux des métazoaires. Le génome mitochondrial est une molécule d'ADN double brin de petite taille (environ 16 kb) généralement circulaire. Des formes linéaires peuvent néanmoins être détectées chez certains organismes (Kayal *et al.*, 2012). Les

deux brins n'ont pas la même densité ce qui permet de distinguer un brin lourd plus riche en guanine (brin H) et un brin léger plus riche en cytosine (brin L). L'ADNmt possède plusieurs particularités qui distinguent sa génétique de celle de l'ADN nucléaire (ADNnu) (tableau 1.1). Entre autres, l'ADNmt est présent en quantité variable; il y a ~ 2 à 10 copies par mitochondries selon le type cellulaire. Une cellule peut donc contenir des milliers voire des centaines de milliers de copies d'ADNmt (Cummins, 1998; Michelle *et al.*, 2001). De fait, le nombre de mitochondries par cellule est lui-même très variable en fonction du type cellulaire (aucune mitochondrie dans les érythrocytes et jusqu'à 100 000 dans les ovocytes chez les mammifères) et peut varier en fonction de l'activité cellulaire (Loublier *et al.*, 2009). À chaque division cellulaire, les mitochondries et les ADNmt se répartissent au hasard dans les cellules filles. Dans une certaine mesure, la génétique mitochondriale est une génétique des populations.

Tableau 1.1 : Comparaison entre les génomes nucléaires et mitochondriaux humains. Tableau modifié de Taylor et Turnbull, (2005).

Caractéristique	Génome nucléaire	Génome mitochondrial
Taille	~3.3 x 10 ⁹ bp	~ 16,500 bp
Nombre de molécules d'ADN par cellule	23 dans les cellules haploïdes; 46 dans les cellules diploïdes	Plusieurs milliers de copies par cellule (polyploïdie)
Quantité absolue dans une cellule	5 pg	0,045 pg
Nombre de gènes encodés	~20,000–30,000	37 (13 polypeptides, 22 ARNt et 2 ARNr)
Densité de gène	~1 par 40,000 bp	1 par 450 bp
Introns	Fréquents dans la plupart des gènes	Absent

Caractéristique	Génome nucléaire	Génome mitochondrial
Pourcentage d'ADN codant	~3%	~93%
Usage des codons	Code génétique universel	AUA code pour la méthionine; TGA code pour le tryptophane; AGA et AGG sont des codons stop.
Protéines associées	Nucléosome-associé à des histones et des protéines non histone	Aucun histones; mais associés à plusieurs protéines (par exemple, TFAM) pour former les nucléoides
Mode d'hérabilité	Hérabilité mendélienne pour les autosomes et le chromosome X; hérabilité paternelle du chromosome Y	Exclusivement maternelle
Réplication	Mécanisme synchrone ; ADN polymérase α et δ	Mécanisme synchrone et asynchrone; ADN polymérase γ seulement
Transcription	La plupart des gènes sont transcrits individuellement	L'ensemble des gènes des 2 brins est transcrit sous forme polycistronique
Recombinaison	Chaque paire de chromosomes homologues recombine durant la prophase de la méiose	Quelques évidences existent au niveau cellulaire, mais très peu d'évidence au niveau des populations

Chez les vertébrés et la plupart des invertébrés, le génome mitochondrial code pour 37 gènes (figure 0.2). Vingt-deux d'entre eux sont des gènes d'ARN de transfert (ARNt) et deux sont des gènes d'ARN ribosomiaux (12S et 16S) codant pour des composantes des dispositifs de transcription de

l'ADNmt elle-même (Ballard et Whitlock, 2004). Les ARNt-mt permettent la lecture des ARNm-mt selon un code génétique propre aux mitochondries. Les treize gènes additionnels codent un ensemble de sous-unités protéiques, toutes des composantes de la chaîne de transport d'électrons : sept sous-unités de la NADH-coenzyme Q oxydoréductase (complexe I ; ND1-6 et ND4L), une sous-unité de la coenzyme Q-cytochrome c oxydoréductase (complexe III ; Cyt b), trois sous-unités de la cytochrome c oxydase (complexe IV ; COI-III) et deux sous-unités de l'ATP-synthase (complexe V; Atpase 6 et 8) (Garesse et Vallejo, 2001). Les autres composantes des complexes enzymatiques du système de transport des électrons et de l'ATPsynthase sont codées par le génome nucléaire. Chaque produit des gènes de l'ADNmt interagit donc avec des protéines codées par ADNnu pour assurer la respiration mitochondriale. La coordination et l'intégration des génomes mitochondriaux et nucléaires représentent une des symbioses les plus fondamentales et remarquables de l'histoire de la vie (Dowling, Friberg et Lindell, 2008).

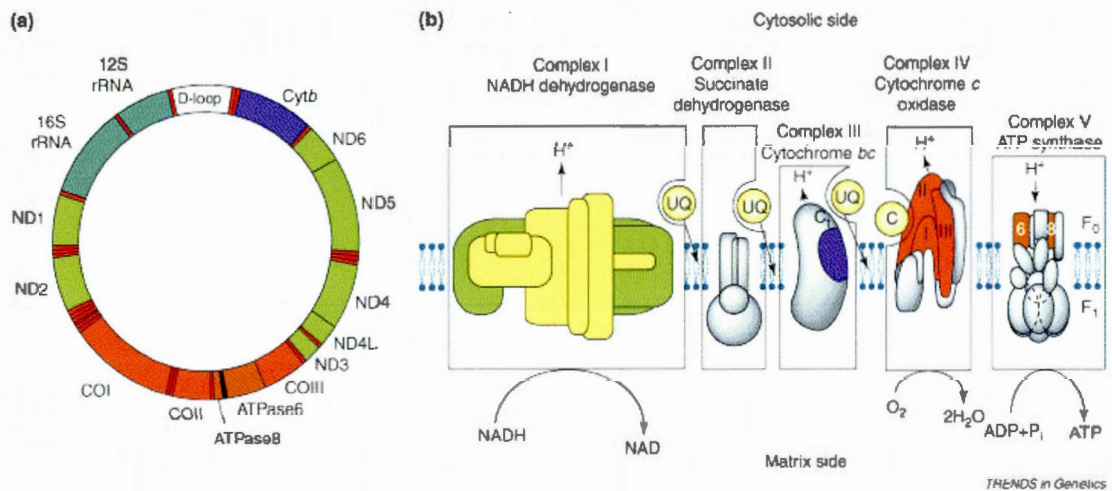


Figure 1.2. (a) ADN mitochondrial humain type codant pour 13 polypeptides entrant dans la composition de (b) la chaîne respiratoire mitochondriale et l'ATP synthase. Tirée de Blier, Dufresne et Burton, (2001).

L'organisation générale du génome mitochondrial chez les animaux témoigne de son évolution progressive vers un très haut niveau de compacité : les deux brins sont codants, il n'y a pas de séquence intergénique, ni d'intron, les gènes sont quasi contigus et quelquefois partiellement chevauchants (Cohen et Lestienne, 2001). Les parties non codantes sont limitées à de très courtes séquences et à une région plus longue appelée la boucle de déplacement ou boucle D (D loop). La boucle D s'étend sur environ 700 paires de bases et contient l'origine de réplication du brin lourd et les origines de transcription des deux brins. La réplication et la transcription de l'ADNmt sont des mécanismes liés dans la mitochondrie. La boucle D est le domaine de régulation du génome mitochondrial, où agissent directement des protéines codées par le génome nucléaire nécessaire à la réplication et à la transcription de cet ADN mitochondrial (Zeviani *et al.*, 1995). L'ADNmt est constamment renouvelée. La fréquence de renouvellement des mitochondries (indépendamment du renouvellement

cellulaire) varie entre dix et 25 jours selon les différents tissus (demi-vie évaluée chez le rat) (Gross, Getz et Rabinowitz, 1969), alors que la fréquence de renouvellement l'ADNmt est seulement de quelques jours (Kai *et al.*, 2006).

1.2.2 Réplication, transcription et traduction de l'ADN mitochondrial

1.2.2.1 Réplication

La réplication, la transcription, la traduction de l'ADNmt sont indépendantes de celles se déroulant dans le noyau et jouent un rôle important relativement au développement des mutations mitochondriales et à son taux d'évolution (Taanman, 1999).

Le réplisome mitochondrial, dispositif multi protéique de réplication, est constituée : (i) d'une ADN polymérase γ (Polg), codée par le génome nucléaire et spécifique à la mitochondrie (Korhonen *et al.*, 2004). Cette enzyme hétérodimère comprend une sous-unité catalytique (PolgA) et une sous-unité accessoire (PolgB) qui se lie à l'ADN, augmentant l'affinité de l'enzyme pour l'ADN et sa processivité (Carrodegua *et al.*, 2001); (ii) la protéine Twinkle, qui a une activité d'hélicase 3'-5' (Spelbrink *et al.*, 2001); (iii) une protéine mtSSB (mitochondrial Single Strand Binding Protein) qui stabilise les structures d'ADN simple brin générées. L'ensemble de ces protéines combinées au facteur de transcription mitochondrial A (TFAM) et aux molécules d'ADNmt forme une structure dynamique appelée nucléoïdes, qui, selon toute vraisemblance serait l'unité de base de la transmission et de l'hérédité mitochondriale (Taylor et Turnbull, 2005).

Deux modèles de réplication de l'ADNmt existent et font encore l'objet de débat dans la littérature (Bogenhagen et Clayton, 2003). Selon le modèle asynchrone et asymétrique (Clayton, 1982; Shadel et Clayton, 1997) la synthèse du brin lourd se produit d'abord et débute au niveau de l'origine de réplication (OH) située dans la boucle D. La réplication du brin lourd se poursuit en exposant au fur et à mesure le brin léger en simple brin. Lorsque l'origine du brin léger OL est atteinte et exposée, environ aux deux tiers du génome, la synthèse de ce brin débute dans la direction opposée. Un deuxième modèle de réplication, celui-là synchrone et symétrique a également été récemment proposé (Holt, Lorimer et Jacobs, 2000; Bowmaker *et al.*, 2003). Ce dernier serait comparable au mécanisme de réplication de l'ADN nucléaire avec une fourche de réplication bidirectionnelle et une synthèse discontinue du brin L par fragments d'Okasaki successifs (Yasukawa *et al.*, 2005).

1.2.2.2 Transcription et traduction

Chaque brin d'ADNmt est transcrit sous forme d'un long ARN polycistronique, à la manière des procaryotes (Clayton, 1991). La transcription de chaque brin est amorcée à partir de deux promoteurs distincts situés dans la boucle D et qui fonctionnent indépendamment et dans des directions opposées. La transcription génère un transcrit primitif qui sera par la suite clivé par des ARNase spécifiques, au niveau des ARN de transfert, libérant ainsi les transcrits des ARNm, ARNr et ARNt. La polyadénylation des ARNm permet ensuite la formation de codon-stop (Fernandez-Silva, Enriquez et Montoya, 2003). Ces ARN sont ensuite traduits par les ribosomes mitochondriaux, suivant un code génétique propre à la mitochondrie (Jukes et Osawa, 1990). Les ribosomes mitochondriaux sont

constitués des ARNr 12S et 16S synthétisés dans la mitochondrie et d'environ 85 protéines importées du cytosol (Fernandez-Silva, Enriquez et Montoya, 2003).

Pour initier la transcription, l'ARN polymérase mitochondriale requiert la présence d'autres facteurs de transcription comme TFAM qui lui permet de se fixer aux zones promotrices des deux brins. Deux autres facteurs régulent spécifiquement l'expression de gènes de protéines mitochondriales. Il s'agit de NRF-1 (nuclear respiratory factor I) et de NRF-2 (nuclear respiratory factor II) (Gleyzer, Vercauteren et Scarpulla, 2005). NRF-1 active la transcription de gènes nucléaires codant pour des protéines des quatre complexes de la chaîne respiratoire et de l'ATPase (Scarpulla, 1997).

1.2.3 Taux de mutation

L'ADNmt aurait un taux de mutation, en moyenne, dix fois supérieur à celui de l'ADN nucléaire (Denver *et al.*, 2000, Haag-Liautard *et al.*, 2008, Lynch *et al.*, 2008). Ce taux varie toutefois de façon importante entre les différents groupes d'organismes (Montooth et Rand, 2008). Le haut niveau de mutagenèse de l'ADNmt peut être attribué à différents facteurs. Entre autres, l'environnement mitochondrial est hautement mutagène et l'ADNmt est localisé à la surface de la membrane de la mitochondrie, côté matriciel, où les espèces oxygénées réactives sont générées au cours de la phosphorylation oxydative (Wallace, 1992). Le risque de dommage oxydatif est d'autant plus important lors de la réplication, puisque l'ADNmt reste longtemps exposé sous forme de simple brin, ce qui le rendrait particulièrement sensible aux attaques radicalaires (Kujoth *et al.*, 2005).

Également, les systèmes de réparation de l'ADNmt sont globalement moins développés et moins efficaces que ceux de l'ADNn (Wilson et Bohr, 2007). Un système mitochondrial de réparation par excision de base BER (base excision repair), impliqué dans la réparation des lésions oxydatives de l'ADN, existe, mais il n'y a aucun système de réparation de type excision de nucléotides NER (nucleotide excision repair) (Croteau et Bohr, 1997; Bogenhagen, 1999). Puisque ce dernier intervient lors de la réplication et que le taux de renouvellement de l'ADNmt est très élevé, l'apparition spontanée de mutation lors de la réplication peut être importante (Bogenhagen, 1999).

Les erreurs de réplication induites par l'ADN polymérase γ peuvent également être un facteur contribuant au taux de mutation élevé de l'ADNmt (Longley *et al.*, 2001). Des mesures *in vitro* ont démontré que cette enzyme était hautement fidèle dans des conditions optimales, mais les biais de nucléotides dans la matrice mitochondriale (*in vivo*) diminueraient son efficacité. Cette polymérase est également responsable d'erreurs de réplication conduisant à des insertions/délétions avec décalage du cadre de lecture au niveau de séquences répétées de l'ADNmt de plus de quatre nucléotides (Longley *et al.*, 2001).

L'absence de protection par des protéines de type histone, protéines que l'on retrouve en association avec l'ADNnu et qui agissent comme barrière mécanique de protection, est également un facteur fréquemment évoqué pour expliquer la mutagenèse de l'ADNmt (Downs et Jackson, 2003).

Le taux de mutation élevé de l'ADNmt rend donc les gènes mitochondriaux susceptibles à une évolution rapide soit via la dérive génétique ou la sélection naturelle et soulève de nombreuses questions

d'ordre évolutives (Birky, Maruyama et Fuerst, 1983; Gemmell, Metcalf et Allendorf, 2004; Bazin, Glemin et Galtier, 2006; Lynch, 2010;).

1.2.4 Homoplasmie et hétéroplasmie

La nature polyploïde du génome mitochondrial – soit le fait que les cellules eucaryotes contiennent chacune un grand nombre de molécules d'ADNmt – donne naissance à une caractéristique singulière de la génétique mitochondriale : le concept d'homoplasmie et d'hétéroplasmie. Lorsque toutes les copies d'ADNmt dans chacune des cellules possèdent une séquence nucléotidique identique, l'individu est dit homoplasmique. La transmission uniparentale couplée au fait que les divisions cellulaires représentent à chaque fois un goulot d'étranglement génétique pour la population mitochondriale contribue à préserver l'homoplasmie (Poulton, 1995). Ainsi théoriquement, un organisme ne possède qu'un haplotype mitochondrial. Toutefois, s'il y a coexistence, dans un même individu, de molécules d'ADN portant des génomes différents du fait d'une mutation ou d'un polymorphisme, l'individu est dit hétéroplasmique.

Le concept d'hétéroplasmie revêt une importance particulière en ce qui a trait aux cytopathies mitochondriales. C'est d'ailleurs dans ce domaine que le plus grand nombre d'études sur le sujet ont été réalisées (Wallace, 1999; Wallace, 2005; Copeland, 2008). Si l'ensemble des copies d'ADNmt est affecté par une mutation, la mutation est dite homoplasmique. Toutefois, dans la mesure où chaque cellule contient une population d'ADNmt, une cellule ou un tissu donné peut contenir un mélange variable de molécules mutées et de molécules sauvages (mutations hétéroplasmiques). En situation d'hétéroplasmie, la sévérité clinique de la pathologie de même que l'impact

sur le fonctionnement du système respiratoire sont déterminés par le pourcentage d'ADNmt muté en relation à l'ADNmt sauvage (Copeland, 2008). Une proportion minimale d'ADNmt mutant est nécessaire pour qu'un tissu présente des dysfonctions métaboliques, phénomène appelé effet de seuil. Puisque les tissus n'ont pas tous la même dépendance au métabolisme oxydatif, leur vulnérabilité aux mutations de l'ADNmt et leur seuil respectif (d'expression de la pathologie) varient d'un tissu à l'autre. Également, au fur et à mesure des divisions cellulaires la proportion d'ADNmt muté et sauvage peut varier en raison du caractère aléatoire de la répartition des ADNmt dans les cellules filles (Wallace, 2005; Copeland, 2008).

La présence de plus d'un variant d'ADNmt (hétéroplasmie) dans un individu est relativement rare et est généralement résolue, c'est-à-dire retourne à un état homoplasmique, en quelques générations (Doublet *et al.*, 2008). Ainsi, dans la plupart des individus, il n'y a pas d'évidence d'hétéroplasmie. Paradoxalement, l'homoplasmie demeure un phénomène conceptuel puisque l'ADN mitochondrial mute constamment; les mutations ponctuelles, les insertions ou les délétions apparaissent et se multiplient de façon clonale ou sont perdues continuellement (Coller *et al.*, 2001; Taylor *et al.*, 2003).

Les mitochondries ne sont pas des organites statiques à l'intérieur de la cellule, elles fusionnent et défusionnent continuellement et forment plutôt un réseau (Zorzano *et al.*, 2010). La fusion et la fission des mitochondries jouent différents rôles tels que le maintien de (i) l'intégrité structurale de la mitochondrie, (ii) la connectivité biochimique et électrique (iii) le renouvellement des mitochondries et (iv) la ségrégation et la protection de l'ADNmt (Turnbull et Lightowlers, 2001; Okamoto et Shaw, 2005). Il a été

également proposé que le processus de fusion et de défusion permettrait la recombinaison des génomes mitochondriaux, un peu à la manière des conjugaisons bactériennes, ce qui contribuerait au maintien de l'homoplasme (Nakada *et al.*, 2001; Ono *et al.*, 2001; Shibata et Ling, 2007). Cette hypothèse allie différentes observations paradoxales qui concernaient la recombinaison des génomes mitochondriaux; par exemple, la présence de la machinerie nécessaire à la recombinaison (Thyagarajan, Padua et Campbell, 1996; Kagawa *et al.*, 2001) et l'apparente absence de recombinaison puisque celles-ci favoriseraient le maintien d'un seul variant mitochondrial rendant la détection de molécule recombinante indétectable. Également, le fait que la recombinaison semble possible au niveau cellulaire, mais jamais entre les populations (Rokas, Ladoukakis et Zouros, 2003; Ladoukakis *et al.*, 2012).

1.2.5 Héritabilité et transmission de l'ADNmt

L'héritabilité du génome mitochondrial est dite non mendélienne, puisque la transmission uniparentale maternelle de l'ADNmt est la règle dans le règne animal (Giles *et al.*, 1980; Birky, 1995). Plus précisément, l'héritabilité uniparentale (quelles soit maternelle ou paternelle) est la caractéristique de la transmission de l'ADN des organites (mitochondries et plastides) commune à un grand nombre de taxons d'eucaryotes sexués, qui inclut les algues, les champignons, les bryophytes (mousses), les gymnospermes (fougères), les plantes supérieures et les animaux incluant l'Homme (Kuroiwa, 2010b). Ainsi chez la plupart des animaux, l'ADNmt dériverait uniquement de la mère. Les individus de sexe mâle possèdent l'ADNmt de leur mère, mais ne le transmettront pas, les mitochondries paternelles et leurs génomes étant spécifiquement éliminés. Ce mode de transmission classique est dénommé héritabilité maternelle stricte ou SMI (strict maternal inheritance).

L'hypothèse principale pour expliquer l'absence de transmission de l'ADNmt paternel fut, pendant longtemps, un simple effet de dilution puisque les gamètes paternels (spermatozoïdes) sont beaucoup plus petits en taille que les gamètes maternels (ovules) et qu'ils ne contribuent quantitativement que très peu à la population mitochondriale transmise à l'embryon (Nishimura *et al.*, 2006). Récemment, cette explication fut remise en question par de nombreuses observations décrivant une digestion active de l'ADNmt des spermatozoïdes, une diminution régulée de la réplication de l'ADNmt des spermatozoïdes et la protéolyse de mitochondries marquées par le système ubiquitine (ubiquitination). Ces observations ont été faites chez le poisson médaka (*Oryzias latipes*), un organisme modèle pour l'étude de la fécondation (Nishimura *et al.*, 2006). Il est maintenant reconnu que l'élimination de l'ADNmt spermatique se produit en deux étapes soit (i) une diminution graduelle du nombre de nucléoides mitochondriaux durant la spermatogénèse et (ii) une digestion rapide de l'ADNmt des spermatozoïdes immédiatement après la fécondation (Nishimura *et al.*, 2006). Plus précisément dans cette dernière étape, les nucléoides mitochondriaux d'ADNmt mâle seraient activement digérés par une nucléase mitochondriale Ca^{2+} -dépendante. La digestion des nucléoides mâles serait ainsi complétée avant la destruction complète des structures mitochondriales, ce qui contribuerait à éviter toute diffusion ou transmission potentiellement délétère de l'ADNmt paternel à la progéniture. Les mitochondries seraient, pour leurs parts, reconnues et détruites par le système ubiquitine-protéasome. Chez les mammifères, ce processus se réalise également en différentes étapes. D'abord, l'ubiquitinylation des mitochondries des spermatozoïdes se déroulerait dans le tractus génital mâle dès le stade de spermatocyte secondaire (Sutovsky *et al.*, 2000). La prohibitine, protéine de la membrane mitochondriale hautement conservée, serait le site marqué par l'ubiquitine

(Thompson, Ramalho-Santos et Sutovsky, 2003). Ensuite, les sites ubiquitinylés seraient masqués par des ponts disulfures au cours du transit épидидymaire et remis à jour uniquement lors de la décondensation du spermatozoïde dans le cytoplasme ovocytaire (May-Panloup *et al.*, 2006). Les mitochondries paternelles seraient par la suite, détruites avant le stade 4 des cellules de l'embryon (Sutovsky *et al.*, 1999).

L'élimination de l'ADNmt paternel de l'embryon n'est pas un mécanisme sans faille. Il a été montré que, dans certaines situations exceptionnelles, l'ADNmt paternel pouvait persister, par exemple dans des embryons humains anormaux (St John *et al.*, 2000) et chez certains individus adultes atteints de myopathies (Vissing et Schwartz, 2002; Williams, 2002).

Il est à noter que le mécanisme de reconnaissance et de destruction des mitochondries paternelles dans le zygote serait spécifique à une espèce donnée. Dans les embryons chimères issus de croisements murins interspécifiques, certaines mitochondries paternelles persistent, alors qu'elles sont absentes chez les embryons issus de croisements d'individus d'une même espèce (Shitara *et al.*, 2000). Des expérimentations réalisées chez la souris ont mis en évidence que ce mécanisme serait également spécifique à la lignée germinale mâle, puisque l'ADNmt des mitochondries provenant de différents tissus et injectées dans un ovocyte de souris n'est pas détruit et peut être détecté chez les nouveau-nés (Shitara *et al.*, 2000).

Les causes évolutives qui ont mené à l'émergence de la transmission uniparentale sont encore loin d'être élucidées. Néanmoins, certaines hypothèses ont été proposées. Il est possible que l'héritabilité uniparentale ait évolué pour limiter la propagation de mutations délétères. Sous un tel mode

d'hérabilité, les mutations apparaissant dans une lignée demeurent dans cette lignée et ne se répandent pas dans la population. Par exemple, l'apparition d'une mutation létale amènerait seulement l'extinction de la lignée et non de la population (Birky, 1995). La deuxième hypothèse, qui n'exclut pas la première, est que l'hérabilité uniparentale serait une stratégie d'évitement des conflits intragénomiques (Cosmides et Tooby, 1981; Burt et Trivers, 2006). Au moment de la fécondation, les mitochondries et leurs ADNmt dérivés de la mère et du père entreraient directement en compétition dans le cytoplasme du zygote. Alors qu'un tel conflit pourrait être coûteux pour le zygote, le génome nucléaire ou un gène nucléaire pouvant supprimer les organites d'un des deux parents et les empêcher de se transmettre représenterait un avantage significatif et se reprendrait dans la population (Xu, 2005; White *et al.*, 2008).

Enfin, une troisième hypothèse propose que durant les premiers stades de l'évolution des eucaryotes, l'hérabilité uniparentale fût un mécanisme crucial dans l'extraction des gènes bactériens symbiotiques du génome de la cellule hôte et dans la formation de la machinerie qui contrôle la division des mitochondries. Kuroiwa *et al.* (1998a) soutiennent que le génome nucléaire de la cellule-hôte a eut à réaliser de nombreuses et complexes étapes (par ex. transfert de près de 80% des gènes, contrôle de la division des protomitochondries) pour en arriver à contrôler la protomitochondrie. L'étape finale de cet asservissement et donc de l'achèvement de l'évolution de la protomitochondrie est la transmission uniparentale. La transmission uniparentale est une façon efficace pour la cellule d'exercer une forme de contrôle sur l'évolution de l'organe en empêchant la recombinaison entre les organites des deux parents et ainsi empêcher une augmentation de la variabilité lors de la reproduction sexuée (Kuroiwa, 2010a).

1.2.6 Évolution de l'ADNmt

Il est aujourd'hui généralement admis que l'ADNmt évolue plus rapidement que l'ADN nucléaire. Initialement et jusqu'à récemment, cette proposition était basée sur le haut degré de divergence observé dans les séquences d'ADNmt entre les espèces. L'ADNmt a d'ailleurs été largement utilisée pour l'étude de la génétique des populations par les biologistes de l'évolution, ces derniers le considérant comme un marqueur génétique neutre (Aise, 1994). Toutefois, les mutations de l'ADNmt étant devenues un sujet central dans l'étude de pathologies humaines, la compréhension de ce phénomène se complexifie et s'accélère. À l'échelle moléculaire, certaines données indiquent un possible rôle de la sélection positive dans l'évolution de l'ADNmt (Blier, Dufresne et Burton, 2001; Mishmar *et al.*, 2003; Schmidt, McCracken et Ferguson-Miller, 2003; Grossman *et al.*, 2004; Ruiz-Pesini *et al.*, 2004; Bazin, Glemin et Galtier, 2006; Meiklejohn, Montooth et Rand, 2007). De plus, de nombreuses études expérimentales confirment la valeur adaptative des variations génétiques qui existent naturellement à l'intérieur et entre les populations (Dowling, Friberg et Lindell, 2008). Il est de plus en plus évident que l'ADNmt serait tout sauf un marqueur neutre.

Une des difficultés est l'inférence du taux de mutation à partir de la divergence des séquences entre espèces. Dans cette situation, ce sont principalement les mutations neutres c'est-à-dire celles n'ayant aucun impact néfaste sur la reproduction ou la survie des individus qui sont détectées (Montooth et Rand, 2008). Il s'agit donc d'une estimation du taux de substitution et non du taux de mutations. Différentes études furent conduites pour évaluer le taux de mutation réel en éliminant l'effet de la sélection naturelle (Parsons *et al.*, 1997; Denver *et al.*, 2000; Rand, 2008; Stewart *et*

al., 2008). De ces dernières, il apparaît que le taux de mutation est environ dix fois supérieur au taux de substitution et que la majorité des mutations ponctuelles est non synonyme et implique un changement d'acides aminés. Ces résultats contrastent avec les hypothèses antérieures et mettent en évidence le fort impact de la sélection naturelle sur le génome mitochondrial. Le rôle de la sélection purificatrice forte éliminant les mutations du génome mitochondrial est maintenant établi et a été observé chez divers taxa (Haag-Liautard *et al.*, 2008; Stewart *et al.*, 2008a, 2008b).

Un certain nombre d'études récentes ont mis en cause l'asexualité du génome mitochondrial comme facteur responsable de la grande charge mutationnelle de ce génome [revue de synthèse par Charlesworth et Wright (2001)]. Il est généralement supposé que dans le génome mitochondrial et dans les génomes nucléaires asexués, le taux élevé de mutation est attribuable à l'absence de recombinaison. Par contre, selon Neiman et Taylor (2009), un très petit nombre d'évènements de recombinaison pourrait contrer l'accumulation de mutation et il a été démontré que le génome mitochondrial se recombine occasionnellement. Une autre possibilité serait que le haut taux d'accumulation de mutations légèrement délétères de l'ADNmt résulterait de la petite taille effective des populations (N_e) associée à sa transmission uniparentale (Lynch, 1997). L'ADNmt subit un intense goulot d'étranglement dans le processus de la transmission uniparentale. Il y aurait, à la fois, un goulot d'étranglement physique de l'ADNmt dans les cellules germinales et un goulot d'étranglement génétique au cours de la folliculogenèse postnatale (Wai, Teoli et Shoubridge, 2008). Ces goulots d'étranglement rendraient le génome mitochondrial effectivement haploïde réduisant ainsi la taille effective de sa population.

Toutefois, l'impact des goulots d'étranglement mitochondriaux et d'une taille effective réduite sur la charge mutationnelle de l'ADNmt demeure une question non résolue. De fait, ces derniers contribuent-ils à réduire (via une augmentation de la variabilité et une intensification de la sélection) ou à augmenter (via la dérive génétique) la charge mutationnelle?

Bazin, Glemin et Galtier (2006) ont montré que le génome mitochondrial diffère du génome nucléaire aussi par le fait que le polymorphisme et le taux d'évolution du génome mitochondrial semblent relativement insensibles aux variations de tailles de population recensée en nature. Afin de comprendre ce qui influence la diversité génétique, Bazin, Glémin et Galtier (2006) analysèrent le polymorphisme de trois bases de données chez environ 3000 espèces animales ; une d'entre elles était constituée de gènes nucléaires, une autre de gènes mitochondriaux et une de données d'allozymes. Alors que les données des gènes nucléaires et des allozymes allaient dans le sens des prédictions des théories de génétique de populations, le polymorphisme de l'ADN mitochondrial ne reflétait pas la taille des populations. Puisque ni le taux de mutation, ni les effets démographiques ne pouvaient expliquer les résultats de leur analyse, ils conclurent que la sélection naturelle (sélection purificatrice ou adaptative) serait en cause. Toutefois, sous un régime de sélection purificatrice, une augmentation de la diversité en fonction de la taille des populations devrait être observée. Puisque c'est une diminution de la variation qui fut observée, ils conclurent que l'ADNmt serait sujet à de fréquents épisodes de sélection naturelle. Plus spécifiquement, ils évoquèrent le processus du « courant génétique » (de l'anglais genetic draft) pour expliquer leurs observations. L'interprétation de ces résultats fut critiquée dans la littérature, certains auteurs soutenant qu'ils reflètent de la sélection

positive et d'autres privilégiant la thèse de la sélection purificatrice [résumé des différentes interprétations dans Meiklejohn, Montooth et Rand, (2007)].

Au-delà des complications propres à la génétique mitochondriale, il faut ajouter que la sélection positive est, de façon générale, très difficile à détecter. Elle opère habituellement de façon épisodique (dans certaines branches d'un arbre phylogénétique) et est souvent restreinte à seulement quelques sites d'acide aminé (Zhang, Nielsen et Yang, 2005). Le signal peut donc être masqué par la sélection négative. Après une trentaine d'années d'étude de l'évolution de l'ADNmt, la génétique de cet organe demeure controversée, mais le développement de nouvelles méthodes bio-informatiques de détection de la sélection positive pourrait permettre de faire la lumière sur l'action de la sélection positive sur le génome mitochondrial [Anisimova et Liberles (2007) pour une revue des différentes méthodes].

1.3 Double transmission uniparentale

1.3.1 L'exception qui confirme la règle

Il existe un système mitochondrial qui transgresse singulièrement les lois de la transmission et de la génétique de l'ADNmt, il s'agit de la double transmission uniparentale [DUI pour « doubly uniparental inheritance » (Zouros *et al.*, 1994)]. Ce système contraste avec l'héritabilité maternelle stricte (SMI) et universelle retrouvée partout ailleurs dans le règne animal. L'existence de la DUI a été démontrée chez certaines familles de mollusques bivalves, dont les moules marines (Mytilidae), les palourdes marines (Veneridae) et les moules d'eau douce (Unionoidae). Les espèces possédant ce système atypique sont caractérisées par la présence de deux ADNmt distincts qui sont hérités respectivement de façon maternelle (ADNmt F) et

paternelle (ADNmt M) (Breton et Beaupré, 2007; Breton *et al.*, 2007). Les divergences observées entre les ADNmt F et M chez les bivalves possédant la DUI peuvent atteindre plus de 10 à 50 % (Doucet-Beaupré *et al.*, 2010). Ces valeurs sont impressionnantes puisqu'il suffit qu'un ADNmt muté diverge d'un seul nucléotide de l'ADNmt natif pour que se développent des pathologies chez l'être humain. Une revue de littérature détaillée du système est présentée dans le chapitre II de la présente thèse. Cette revue, bien que publiée en 2007, demeure dans l'ensemble pertinente bien que des développements récents sont venus la compléter. Ces avancements, de même que quelques aspects fondamentaux du système, sont donc décrits ci-dessous.

1.3.2 DUI : le modèle

Dans le système de la DUI, deux ADNmt distincts coexistent, mais leur distribution dans les différents tissus des individus varie. Typiquement, les femelles sont homoplasmiques et contiennent uniquement l'ADNmt F tandis que les mâles sont hétéroplasmiques [revue dans Breton *et al.* (2007) et Passamonti et Ghiselli (2009)]. Chez ces derniers, l'ADNmt F domine dans les tissus somatiques alors que la gonade contient presque exclusivement l'ADNmt M. Les deux sexes héritent de l'ADNmt F alors que l'ADNmt M n'est transmis que de père en fils. Conséquemment, la DUI n'est pas de la transmission biparentale, chaque variant mitochondriaux se transmettant de façon parallèle, mais uniparentale.

Plus explicitement, les mitochondries des gamètes mâles et femelles forment la population mitochondriale dans l'œuf fertilisé. Le modèle actuel de la DUI prédit que le génotype nucléaire de l'œuf est responsable du

comportement des mitochondries : les œufs possédant l'allèle nucléaire dominant Z vont produire le facteur Z. Ce facteur, dont la nature exacte demeure encore inconnue, bloque un système d'élimination des mitochondries paternelles similaire à l'ubiquitine et permet la rétention et l'agrégation des mitochondries paternelles qui seront préférentiellement canalisées vers les cellules primordiales qui formeront la gonade, rendant ainsi l'ADNmt paternel prédominant chez la lignée germinale mâle (Kenchington *et al.*, 2002). Les œufs récessifs zz ne produisent pas de facteur Z et les mitochondries paternelles se dispersent aléatoirement dans les cellules et/ou sont détruites dans les embryons destinés à devenir des femelles.

Il faut toutefois ajouter quelques nuances à ce modèle. Par exemple, le mécanisme de ségrégation de l'ADNmt M n'est pas toujours parfait (Zouros, 2000) : des traces de génome M peuvent être trouvées dans les tissus somatiques femelles (Garrido-Ramos *et al.*, 1998; Dalziel et Stewart, 2002), de même que dans les œufs non fertilisés de *Mytilus galloprovincialis* (Obata *et al.*, 2006). Par contre, la lignée germinale mâle semble exclure systématiquement le génome F (Venetis *et al.*, 2006).

1.3.3. Prévalence de la DUI chez les bivalves

Le système de la DUI a été initialement découvert dans les années 1990 chez des populations de moules marines du genre *Mytilus* au sud-ouest de l'Angleterre et l'étude de ce système s'est poursuivie principalement chez ce groupe (Fisher et Skibinski, 1990). Par la suite, le système a été confirmé chez deux autres familles de bivalves soit les Unionidae et les Veneridae (Rawson et Hilbish, 1995; Stewart *et al.*, 1995; Hoeh *et al.*, 1996; Stewart *et*

al., 1996; Passamonti et Scali, 2001; Curole et Kocher, 2002; Passamonti, Boore et Scali, 2003; Walker *et al.*, 2006). Les études visant la détection de la DUI demeurent toutefois difficiles et marginales. Il n'est donc pas possible d'affirmer que la distribution réelle de la DUI est totalement connue. Pourtant une connaissance approfondie de l'étendue de la distribution taxonomique de la DUI permettrait de mieux s'attaquer à la question de l'origine unique par rapport à l'apparition multiple (dans diverses lignées éloignées) de la DUI. Quelques études récentes ont tout de même permis de confirmer ou d'infirmer la présence de ce système atypique. Par exemple, l'absence de la DUI a été confirmée chez l'huître (Obata *et al.*, 2008). De plus, sa présence a été confirmée chez une autre espèce de Mytilidae *Musculista senhousia* (Passamonti, 2007). Plus significatif, la DUI a aussi été confirmée chez la coque *Donax trunculus* qui fait partie d'un autre groupe de bivalves les Donacidae (Theologidis *et al.*, 2008). La DUI s'observe donc chez sept familles appartenant à dans trois grandes sous-classes de bivalves soit : (i) chez les ptériomorphes (moules marines) dans la famille des Mytilidae, (ii) chez les hétérodontes (palourdes marines) dans les familles des Donacidae, des Solenidae et des Veneridae) et (iii) chez paléohétérodontes (moules d'eau douce) dans les familles des Unionidae, des Hyriidae et des Margaritiferidae. (Passamonti et Ghiselli, 2009).

1.3.4 Particularités de la DUI dans les différentes familles de bivalves

L'avancement de la recherche sur le système de transmission mitochondriale atypique de la DUI a permis de mettre en évidence des différences évolutives importantes entre les différents groupes de bivalves qui possèdent la DUI. Par exemple chez les moules marines, certains mâles, au lieu de posséder un ADNmt M standard, possèdent un ADN de type F

recombinant (qui possède une origine de réplication mâle, mais une séquence codante maternelle) dans leur gonade, et qui se transmet de père en fils par les spermatozoïdes comme un ADNmt M standard. Ce phénomène est appelé renversement des rôles ou encore phénomène de masculinisation parce qu'un ADNmt de type F a «envahi» la lignée de transmission paternelle et assume maintenant le rôle d'un génome M dans les spermatozoïdes (Hoeh *et al.*, 1997; Quesada, Wenne et Skibinski, 1999).

Breton *et al.* (2006) ont publié un génome complet de *Mytilus trossulus* qu'ils avaient identifié comme un génome récemment masculinisé (présenté comme tel dans la revue de littérature du chapitre II). Toutefois, le fait que ce génome soit un génome récemment masculinisé a été remis en doute par certains auteurs (Venetis *et al.*, 2007; Cao *et al.*, 2009). Selon ces derniers, il s'agirait plutôt du génome F qui se retrouve normalement dans les tissus somatiques chez le mâle. Le séquençage d'un troisième type de génome chez *M. galloprovincialis* a été effectué et son analyse démontre qu'il possède une région de contrôle trois fois plus longue que celle des génomes M et F standards et qui combine à la fois les régions de contrôle des génomes M et F. Selon Cao *et al.* (2009), ce génome représente un «vrai» génome récemment masculinisé.

De nombreux génomes recombinants ont également été trouvés dans des populations européennes de *Mytilus* et contrairement aux observations précédentes certains des génomes recombinants n'étaient pas associés à un sexe en particulier (Filipowicz *et al.*, 2008). Il semble que les génomes recombinants ne suivent pas les règles de transmission de la DUI ou encore peuvent changer de lignée facilement.

L'existence de la masculinisation et de molécules ADNmt recombinantes dans le système de la DUI permet de tester différentes hypothèses physiologiques. Deux études ont utilisé ce modèle pour comparer la vitesse de nage et l'activité des enzymes mitochondriales des spermatozoïdes contenant un ADNmt M versus ceux contenant un ADNmt F recombinant (Jha *et al.*, 2008; Breton, Stewart et Blier, 2009). Les résultats ont montré que les spermatozoïdes ayant de l'ADNmt M standard possédaient une motilité et des activités enzymatiques mitochondriales diminuées. Ces résultats allaient à l'encontre de la prémisse de l'étude selon laquelle l'existence d'un variant mitochondrial paternel pourrait permettre une adaptation de l'ADNmt M pour des fonctions spermatiques.

Jusqu'à présent, aucune évidence de recombinaison entre les génomes M et F ni aucun événement de masculinisation n'a été observé chez les moules d'eau douce (Unionoidae) (Hoeh, Stewart et Guttman, 2002). L'absence de masculinisation coïncide avec une particularité de la DUI retrouvée uniquement chez ce groupe soit la présence d'une extension en 3' sur la cytochrome c oxydase du génome M (Mcox2ext) (Curole et Kocher, 2002; Curole et Kocher, 2005). Cette extension mesure environ 185 codons et a été observée chez tous les Unionoïdés examinés jusqu'à présent. La Mcox2ext serait le domaine mitochondrial qui évolue le plus rapidement (Curole et Kocher, 2005). Elle ne serait pas présente chez les moules marines ni chez aucun autre animal. La Mcox2ext formerait des hélices transmembranaires et des analyses d'immunohystochimie démontrent qu'elle se localiserait dans la membrane interne, mais aussi dans la membrane externe (Chakrabarti *et al.*, 2006).

Récemment, l'étude de cette *Mcox2ext* a permis d'établir qu'elle s'exprimerait de façon saisonnière dans les gonades males, avec un maximum d'expression immédiatement avant la fertilisation et ce, indépendamment de la densité de spermatozoïdes (Chakrabarti *et al.*, 2007). Elle s'exprimerait également dans les gonades femelles. L'hypothèse mise de l'avant est que la *Mcox2ext* jouerait un rôle dans la reproduction soit dans la maturation des gamètes, la fertilisation et/ou l'embryogenèse (Chakrabarti *et al.*, 2009). De plus, la *Mcox2ext* serait fonctionnelle, ses structures primaires et secondaires évoluent rapidement et certains sites spécifiques d'acides aminés seraient possiblement sous sélection positive (Chapman *et al.*, 2008).

1.3.5 Évolution des génomes F et M

La DUI représente un système unique pour examiner l'effet des forces évolutives agissant sur le génome mitochondrial. L'analyse de portions de gènes des génomes M et F chez différentes espèces a indiqué que les génomes M et F évoluaient tous deux plus rapidement que les ADN mitochondriaux des autres taxa animaux et que le génome M évoluait plus rapidement que le génome F (Rawson et Hilbish, 1995; Stewart *et al.*, 1995; Hoeh *et al.*, 1996a; Liu, Mitton et Wu, 1996; Curole et Kocher, 2002; Passamonti, Boore et Scali, 2003; Curole et Kocher, 2005). L'hypothèse principale de cette évolution accélérée est une relaxation des contraintes sélectives attribuable à une division inégale de la charge de travail dans le système de la DUI (Stewart *et al.*, 1996). Contrairement à l'ADNmt typique qui doit fonctionner dans les gonades et les tissus somatiques des deux sexes, le génome F fonctionne que dans la gonade femelle et dans les tissus somatiques M et F alors que le génome M a à fonctionner seulement dans la gonade mâle. Lors d'une étude de génétique des populations chez *Mytilus*

californianus, Ort et Pogson (2007) ont trouvé des traces d'une forte sélection purificatrice étaient observables dans les deux génomes, mais semblait moins intense chez le génome mâle. Plus récemment, dans une étude sur des populations européennes de *Mytilus*, Śmietanka, Burzynski et Wenne (2009) conclurent essentiellement que les patrons généraux de polymorphisme observés concordaient avec (i) une forte sélection purificatrice, (ii) une relaxation des contraintes sélectives sur le génome M comparativement au génome F, mais aussi que (iii) l'évolution du génome M serait plus souvent interrompue par des épisodes de dérive génétique attribuables à des épisodes d'évolution adaptative (courant génétique) [voir la discussion de Śmietanka, Burzynski et Wenne, 2009].

La relaxation des contraintes sélectives sur le génome mâle n'explique ni l'origine ni le maintien du système de la DUI. En effet, le fait qu'un système aussi atypique que celui de la DUI se maintienne pourrait laisser croire qu'il y a une raison adaptative sous-jacente et que cela n'est pas un simple état de fait sans quoi ce dernier aurait probablement été perdu.

La duplication de gènes ou de génomes complets dans un système est reconnue comme un phénomène important au cours de l'évolution des eucaryotes. La plupart des gènes dupliqués expérimenteraient une période de relaxation des contraintes sélectives (favorisant un taux de mutation élevé). Puis, les gènes ou les génomes dupliqués seraient destinés soit à devenir non fonctionnels et/ou disparaître, soit à développer une nouvelle fonction via le processus de la subfonctionnalisation ou la néofonctionnalisation et être ainsi retenus dans le génome ou dans le système (Ohno, 1970; Nowak *et al.*, 1997; Huminiecki et Wolfe, 2004; Roth *et al.*, 2007).

Très tôt, Skibinski, Gallagher et Beynon (1994a, 1994b) avaient suggéré que la DUI avait évolué pour permettre l'évolution adaptative sur le génome mâle. Une des considérations évolutives fascinantes de la DUI, est qu'elle pourrait peut-être permettre à la sélection d'agir directement sur l'ADNmt paternel puisqu'il ne représente plus un cul-de-sac évolutif comme c'est le cas dans la transmission maternelle. Jusqu'à maintenant, quelques indices laissant croire que la sélection positive pourrait agir dans le système de la DUI se sont accumulés, que ce soit dans les études de génétiques des populations ou au niveau de l'évolution moléculaire de la *Mcox2ext* par exemple, mais un tel processus n'a toujours pas été démontré.

1.3.6 Objectif général du projet de recherche

L'objectif général du présent projet de recherche est de mettre en exergue le rôle des contraintes sélectives sur l'évolution moléculaire des génomes mitochondriaux mâle et femelle du système de la double transmission uniparentale chez les Unionidés. Chez ce groupe, la double transmission uniparentale semble évolutivement plus stable (aucun évènement de masculinisation n'a été recensé) et la divergence entre les ADNmt M et F est significativement plus élevée que chez les taxa marins. La double transmission uniparentale représente, certes, un système atypique; mais les descriptions, les comparaisons et les analyses qu'elle permet contribuent à développer une image plus exacte de l'évolution, et cela, non seulement au niveau de l'évolution de la DUI mais bien de l'évolution du génome mitochondrial dans son ensemble.

1.3.7 Objectifs spécifiques du projet de recherche

Chapitre II

Tel que précisé précédemment, le Chapitre II est une revue de littérature sur le système de la double transmission uniparentale. Les observations qui ont menées à la découverte de la DUI y sont présentées de même que le modèle et les mécanismes de la DUI. Cette revue met en évidence que, bien que l'essentiel de la structure du système de la DUI soit connu, les implications évolutives de cette transmission atypique demeurent énigmatiques et les hypothèses qui lui sont associées restent à peaufiner. Deux hypothèses sur l'origine et le maintien du système de la DUI chez certains bivalves au cours de l'histoire évolutive y sont présentées. Également, cette revue avait comme objectif de proposer la DUI comme un système modèle qui pourrait nous permettre d'acquérir de nouvelles connaissances sur les interactions et la coadaptation des génomes nucléaires et mitochondriaux. En effet, ce système offre la possibilité d'étudier deux variants mitochondriaux fortement divergents qui coexistent et s'adaptent simultanément à un génome nucléaire.

Chapitre III

L'objectif de ce premier chapitre de recherche est de séquencer, d'annoter et de publier pour la toute première fois des génomes mâles complets d'Unionoïdés. La réalisation de cet objectif fournit un contexte pour les comparaisons des génomes mitochondriaux chez les lignées de bivalves possédant la DUI et les lignées ne la possédant pas. De ce fait, nous

pouvons vérifier l'hypothèse voulant que les Unionoïdés possèdent le caractère ancestral de la DUI. Enfin, cette étude a également comme objectif d'identifier les caractéristiques putatives des génomes mitochondriaux mâles impliquées dans la manifestation de la DUI. Dans la présente étude, six nouveaux génomes mitochondriaux complets d'espèces d'Unionoïdés, soit le génome F et M de *Venustaconcha ellipsiformis* (Unionoida: Unionidae: Ambleminae: Lampsilini), *Pyganodon grandis* (Unionoida: Unionidae: Unioninae: Anodontini) et *Quadrula quadrula* (Unionoida: Unionidae: Ambleminae: Quadrulini), ont été comparés avec les génomes complets disponibles d'espèces possédant la DUI déposées dans GenBank. Ces descriptions et comparaisons ont permis de dresser un portrait plus complet de l'évolution du système de la DUI.

Chapitre IV

Dans ce chapitre, l'objectif est d'analyser l'évolution des génomes M et F dans un cadre phylogénétique défini et plus restreint. Le genre *Pyganodon* a été choisi en raison de sa richesse taxonomique, de sa vaste répartition géographique dans le nord-est de l'Amérique du Nord et de la disponibilité de génomes M et F complets (chapitre III). L'exploration de la distribution de la variabilité génétique entre et à l'intérieur des populations et des espèces de *Pyganodon* a permis la comparaison des processus évolutifs impliqués entre les génomes mitochondriaux F et M chez les Unionidés. Les objectifs spécifiques de ce chapitre sont de (1) clarifier les relations phylogénétiques à l'intérieur du genre *Pyganodon* (Unionidae, Bivalvia) dans le nord-est de l'Amérique du Nord en utilisant des séquences (cox1 et cox2) des génomes mitochondriaux F et M ; (2) comparer et confronter les signaux phylogénétiques rendus par les génomes M et F à l'intérieur de ce groupe; (3)

et d'analyser selon le contexte géographique et historique les patrons de variation génétique observés à l'intérieur et entre les lignées.

Chapitre V

L'objectif de ce dernier chapitre de recherche était de vérifier si les génomes M et/ou les génomes F des groupes possédant la DUI sont soumis à un régime de sélection positive. Plus spécifiquement, nous avons utilisé un test de sélection en maximum de vraisemblance de type branche-site afin de déterminer si certaines branches des arbres phylogénétiques des bivalves (chapitre III) étaient soumises à de la sélection positive. Les génomes se retrouvant à l'intérieur des branches où la sélection positive fut détectée de façon significative furent soumis à un test BEB (Bayes Empirical Bayes) *a posteriori* afin de déterminer quels étaient les sites potentiels où la sélection positive opérait.

CHAPITRE II

THE UNUSUAL SYSTEM OF DOUBLY UNIPARENTAL INHERITANCE OF MTDNA: ISN'T ONE ENOUGH?

Sophie Breton¹, *, Hélène Doucet-Beaupré¹, *, Donald T. Stewart², Walter R. Hoeh³ and Pierre U. Blier¹

¹ Laboratoire de Biologie Intégrative, Département de Biologie, Université du Québec à Rimouski, Rimouski, G5L 3A1 Canada ; (H. Doucet-Beaupré et S. Breton: rédaction de l'article et conception des figures ; P. U. Blier : aide à la rédaction).

² Department of Biology, Acadia University, 24 University Avenue, Wolfville, Nova Scotia, B4P 2R6, Canada; (D. T. Stewart: aide à la rédaction).

³ Department of Biological Sciences, Kent State University, Kent, Ohio 44242, USA; (W. R. Hoeh: aide à la rédaction).

* These authors contributed equally to this work.

Paru en 2007 dans *Trends in Genetics*, vol. 23, no 9, p. 465-474.

Résumé

Les mitochondries possèdent leur propre matériel génétique, l'ADN mitochondrial (ADNmt). Chez les vertébrés et la plupart des invertébrés, l'ADNmt code pour treize sous-unités protéiques entrant dans la composition des complexes enzymatiques du système de transport des électrons et de l'ATPsynthase de même que pour des composantes impliquées dans la transcription et la traduction de l'ADNmt lui-même. Dans le règne animal, l'ADNmt est transmis exclusivement de façon maternelle (SMI pour *strict maternal inheritance*). La découverte de la «double transmission uniparentale (DUI)» de l'ADNmt chez certains groupes de bivalves interroge les lois de transmission des mitochondries chez les animaux. Dans ce chapitre, nous passons en revue les avancées récentes dans notre compréhension de la DUI. Les espèces possédant ce système particulier sont caractérisées par la présence de deux ADNmt distincts qui sont hérités de façon maternelle (ADNmt F) ou paternelle (ADNmt M). Les particularités de ce système entre les différents groupes de bivalves y sont présentées entre autres la présence d'une nouvelle extension d'un gène mitochondrial, la recombinaison et les «renversements de rôle» périodiques des génomes mitochondriaux mâle et femelle. Le système de la DUI offre une occasion unique d'étudier les interactions nucléomitochondriales et la signification évolutive de différents modes d'hérédité mitochondriale.

2.1 Abstract

Mitochondria possess their own genetic material (mitochondrial DNA or mtDNA), which of that gene products are involved in mitochondrial respiration and oxidative phosphorylation, transcription, and translation. In animals, mitochondrial DNA is typically transmitted to offspring by the mother alone. The discovery of 'doubly uniparental inheritance' (DUI) of mtDNA in some bivalves has challenged the paradigm of strict maternal inheritance (SMI). In this review, we survey recent advances in our understanding of DUI, which is a peculiar system of cytoplasmic DNA inheritance that involves distinct maternal and paternal routes of mtDNA transmission, a novel extension of a mitochondrial gene (*cox2*), recombination, and periodic 'role-reversals' of the normally male and female-transmitted mitochondrial genomes. DUI provides a unique opportunity for studying nuclear-cytoplasmic genome interactions and the evolutionary significance of different modes of mitochondrial inheritance.

2.2 Mitochondrial inheritance: rules for mussels

Mitochondria are multifunctional, DNA-bearing organelles found in eukaryotic cells. Animal mitochondrial DNA (mtDNA) is typically a circular molecule ~16.5-kilobase long [but linear and longer mtDNAs exist across eukaryotes, see Ref. Burger, Gray and Lang, 2003] that normally encodes ~37 genes (Boore, 1999). Among them, 24 mitochondrial genes encode components involved in the mitochondrial translational machinery (22 tRNAs and two rRNAs). The remaining 13 genes encode protein subunits of the respiratory chain complexes and ATP synthase (see Glossary).

In animal species, mtDNA is exclusively maternally inherited (Birky, 2001) (but see Refs (Xu, 2005) and (Elson and Lightowlers, 2006) for exceptions). Apart from rare mutants, all copies of mtDNA in each cell typically have identical DNA sequences, a situation known as homoplasmy. Disruption of mitochondrial homoplasmy (i.e. the mixing of different mtDNAs within a cell or heteroplasmy) results in genetic variance within an organism. The resulting potential for inter-mitochondrial competition could set the stage for the spread of deleterious, selfish mitochondrial elements (Wallace, 2005) and (Burt and Trivers, 2006). Uniparental inheritance of mtDNA in animals is a mechanism that has apparently evolved to avoid such intracellular conflict (Rand, 2001; Ballard and Whitlock, 2004; Burt and Trivers, 2006). However, an extreme deviation from this general rule, termed doubly uniparental inheritance (DUI) (Zouros *et al.*, 1994b), occurs in marinemussels of the order Mytiloida, freshwater mussels of the superfamily Unionoidea, and marine clams of the order Veneroida (Skibinski, Gallagher and Beynon, 1994a; 1994b; Zouros *et al.*, 1994a; 1994b; Hoeh *et al.*, 1996b; Liu, Mitton and Wu, 1996; Passamonti and Scali, 2001; Curole and Kocher, 2002; Hoeh, Stewart and Guttman, 2002; Passamonti, Boore and Scali, 2003; Curole and Kocher, 2005; Walker

et al., 2006). Species possessing DUI are characterized by the presence of distinct gender-associated mtDNAs that are inherited either maternally or paternally. These female-transmitted and male-transmitted mitochondrial genomes (referred to for convenience as 'F genomes' and 'M genomes', respectively) often exhibit nucleotide divergences greater than 20%. Female mussels typically inherit their F genome only from their mother, but they transmit this F genome to both sons and daughters. Male zygotes inherit their mtDNA from both parents, but they sort the mixture of mitochondrial genomes present such that the M genome inherited from their father becomes established in the germ line. This paternal M genome will subsequently be passed on by way of the sperm where it will ultimately be retained only by sons, possibly as a consequence of limited replication of the M genome in females (Obata *et al.*, 2006; Sano, Obata and Komaru, 2007). Females are essentially homoplasmic for the F genome whereas males are heteroplasmic for both F and M genomes. This peculiar situation challenges our traditional view of the strict maternal inheritance (SMI) of mtDNA. A striking difference between SMI and DUI systems is that the latter allows selection to act directly on the male mitochondrial genome; unlike most animals, male mussels do not represent an evolutionary dead-end for mitochondrial genomes (Zeh, 2004; Zeh and Zeh, 2005).

Although the essential structure of the DUI system is known, we are still revising and refining our understanding of the evolutionary implications of DUI. Major questions about DUI into which we are just now gaining insight include whether DUI originated to avoid the deleterious effects of mitochondrial DNA mutations on sperm function or whether the phenomenon evolved because of a role in sex determination of selfish mitochondrial DNA elements. In addition to surveying recent studies of DUI, we discuss the latest

ideas regarding the evolutionary origin of this atypical system of mtDNA inheritance and we highlight some fascinating questions that are beginning to emerge from studies of DUI.

2.3 The DUI system unveiled

DUI was discovered in 1990 when a high frequency of heteroplasmic individuals was detected in a study of mtDNA variation in a hybrid zone between *Mytilus edulis* and *M. galloprovincialis* mussels in southwest England (Fisher and Skibinski, 1990). The occurrence of two divergent mtDNAs in the same individuals was subsequently confirmed in other mytilid populations (Hoeh, Blakley and Brown, 1991; Zouros *et al.*, 1992). These findings, combined with previous cytological studies showing retention of paternal mitochondria in early embryos (Longo and Anderson, 1969), led the authors to suggest biparental transmission of *Mytilus* mitochondrial DNA (Hoeh, Blakley and Brown, 1991; Zouros *et al.*, 1992). Subsequent studies showed that heteroplasmy in *Mytilus* were associated to a gender-associated mtDNA transmission system that required distinct paternal and maternal mitotypes (Zouros *et al.*, 1994a; 1994b). For this reason, it was named doubly uniparental inheritance (Zouros *et al.*, 1994a). DUI was later documented in other (but not all) bivalve taxa as noted above (Figure 2.1). Given the broad taxonomic distribution of DUI observed within the Bivalvia (Rawson and Hilbish, 1995; Stewart *et al.*, 1995; Hoeh *et al.*, 1996b; Stewart *et al.*, 1996; Passamonti and Scali, 2001; Curole and Kocher, 2002; Passamonti, Boore and Scali, 2003; Walker *et al.*, 2006) it was proposed that this system evolved once in an ancestral bivalve lineage and was subsequently lost in some descendants (Zouros, 2000; Hoeh, Stewart and Guttman, 2002).

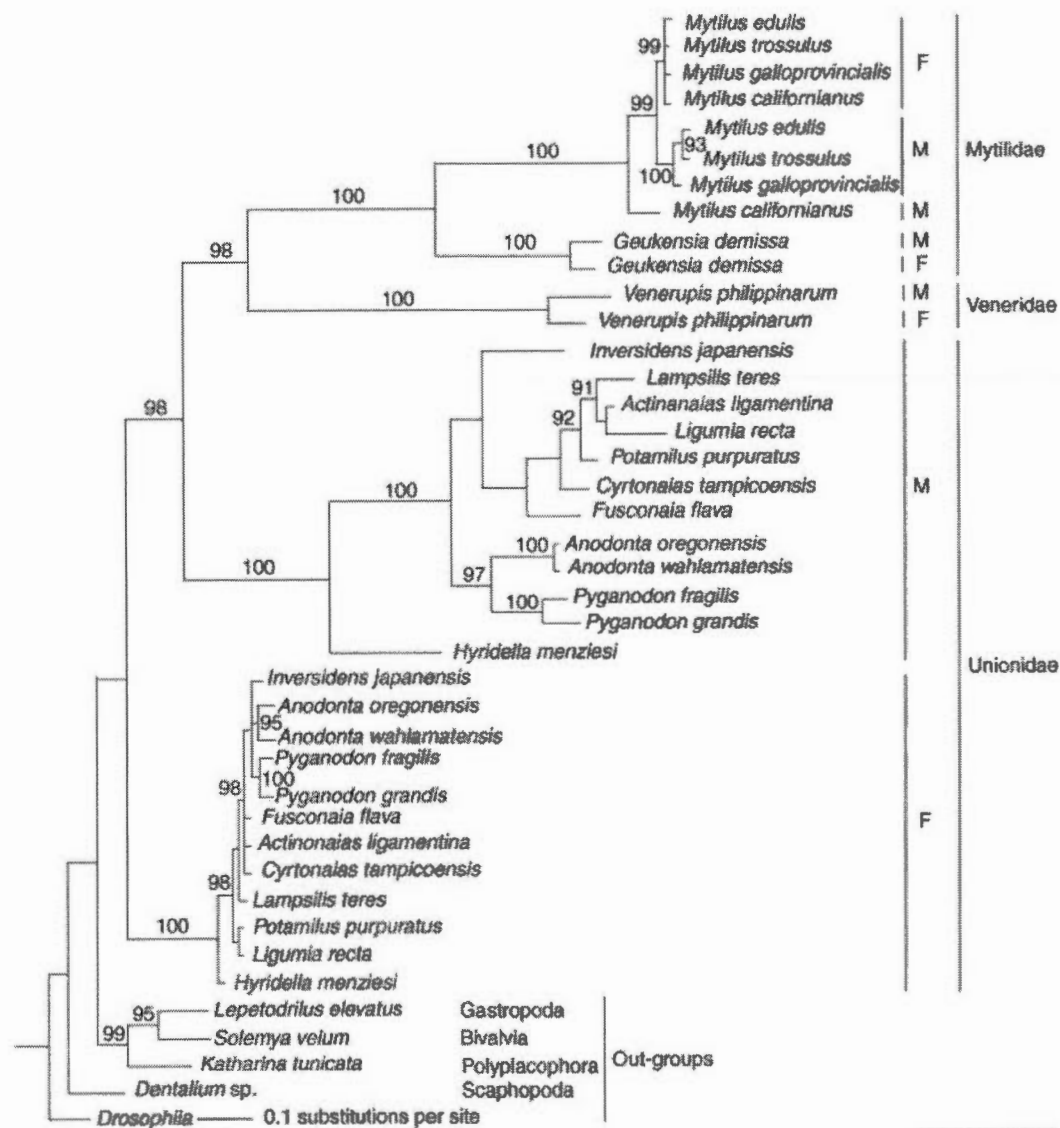


Figure 2.1 DUI phylogeny. The mitochondrial genome evolutionary relationships for the DUI-containing taxa showing that the male (M) and female (F) genomes in the freshwater mussels, the Unionidae, have been stable, whereas there is evidence for occasional masculinization events (i.e. F genomes giving rise to M genomes) in the families Veneridae and Mytilidae. Topology and branch lengths are based on analysis of 199 amino acids from *cox1* and posterior probabilities ($\times 100$) > 90 are displayed. The taxonomic and

sex-specific transmission affiliations of the individual sequences are indicated on the right. The Mr. Bayes analysis (Huelsenbeck and Ronquist, 2001) used the Jones, Taylor and Thornton (JTT) model of amino acid substitution (Jones, Taylor and Thornton, 1992) and was run for 10 million generations with eight chains. Overall, 10 000 trees were saved during the course of the analysis but only the last 9000 were used to construct the consensus tree (i.e. the first 1000 trees were discarded as burn-in). The original nucleotide sequences (species and accessions numbers) were translated with the *Drosophila* mitochondrial genetic code and used according to Refs (Beagley, Taylor and Wolstenholme, 1997; Hoeh *et al.*, 1998; Hoeh, Stewart and Guttman, 2002; Boore, Medina and Rosenberg, 2004; Mock *et al.*, 2004; Riginos *et al.*, 2004; Mizi *et al.*, 2005; Breton *et al.*, 2006).

The transition from strict maternal inheritance (SMI) to DUI probably involved a modification of the recognition system of sperm mitochondria by eggs, and a specific mechanism ensuring a father-to-son transmission of M mtDNA (Zouros, 2000). Empirical evidence for such mechanisms came from studies on *Mytilus* embryos. Typically, the fate of sperm mitochondria depends on whether the embryo is destined to develop into a female or a male (Saavedra, Reyero and Zouros, 1997; Sutherland *et al.*, 1998; Cao, Kenchington and Zouros, 2004; Cogswell, Kenchington and Zouros, 2006; Venetis *et al.*, 2006). For example, Sutherland *et al.* (1998) found that during fertilization, all eggs receive sperm mitochondria, which are eliminated or drastically reduced within 24–48 hours in female embryos. Recent epifluorescence-based observations of embryos destined to become males demonstrated that sperm mitochondria tend to aggregate in a single blastomere that is thought to give rise to the male germ line. By contrast, in embryos destined to become female, sperm mitochondria are randomly dispersed among blastomeres (Cao, Kenchington and Zouros, 2004; Cogswell, Kenchington and Zouros, 2006). This sex-specific difference in the embryonic aggregation of M versus F genomes, which is one factor responsible for the tissue-specific differences in ratios of M and F genomes,

appears to be dependent on the action of microtubules (Yaffe, 1999; Obata and Komaru, 2005).

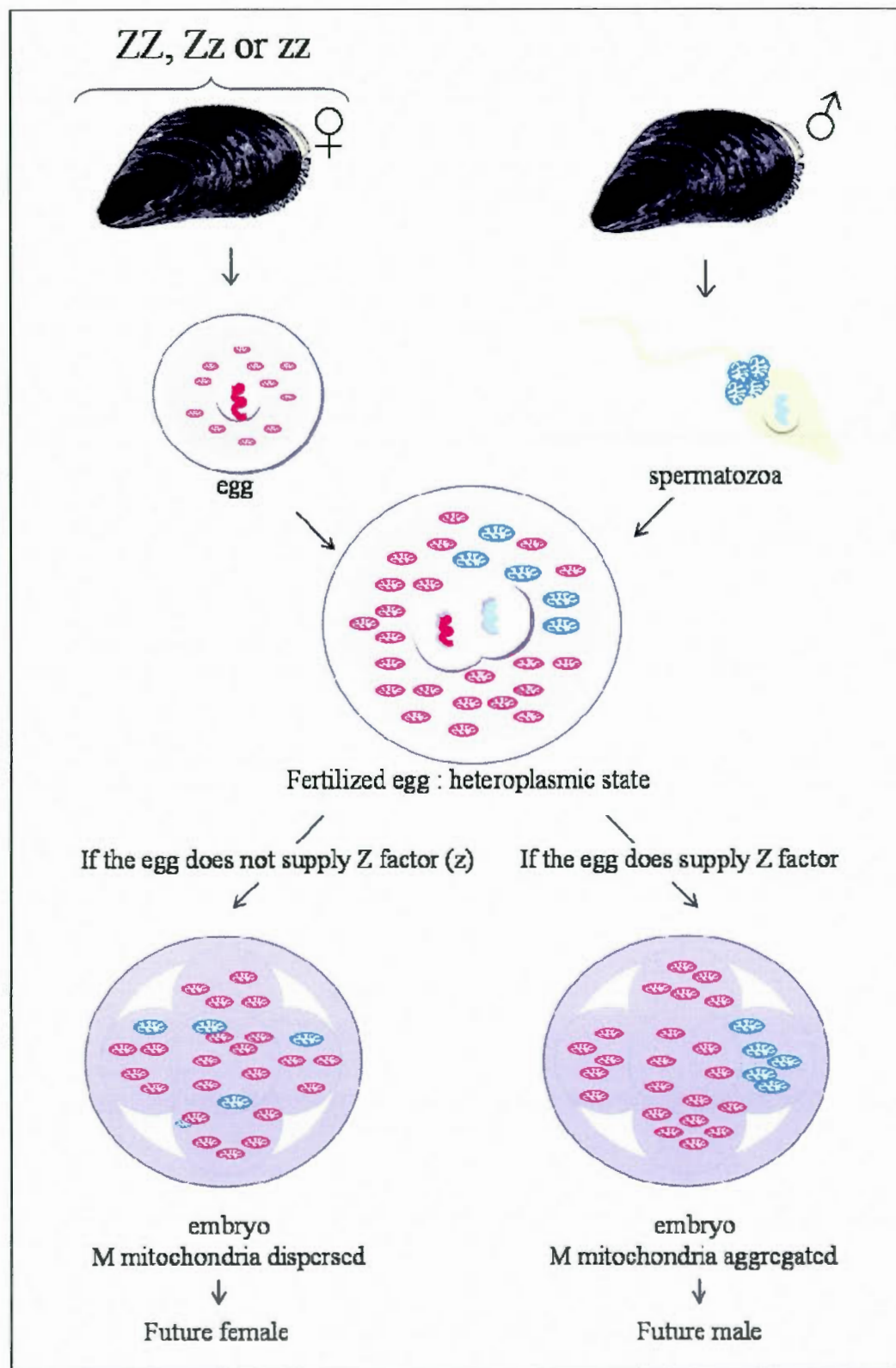
These studies suggest that heteroplasmy is the initial state in the early development of mytilid embryos. In the family Mytilidae, females normally shift from being heteroplasmic zygotes to essentially homoplasmic (F genome containing) adults. Mature male mytilids contain varying ratios of the F and M genomes in all tissues (i.e. testes contain predominantly M genome, somatic tissues contain predominantly F genome) (Garrido-Ramos *et al.*, 1998; Dalziel and Stewart, 2002). By contrast, the venerid clam *Venerupis* (= *Ruditapes*) *philippinarum* has a strong predominance of M mtDNAs in somatic tissues (Passamonti and Scali, 2001). Taken together, these results indicate that mechanisms for the sequestration of sperm mitochondria and the M genome into the male gonad are not perfect (Zouros, 2000), or that M genomes have distinct functional repertoires in these two divergent marine taxa. Similarly, the mechanisms that limit the presence of sperm mitochondria and M genomes in developing females are also not perfect. Trace amounts of M genome have been found in tissues of adult females (Garrido-Ramos *et al.*, 1998; Dalziel and Stewart, 2002) and even in the unfertilized eggs of *M. galloprovincialis* (Obata *et al.*, 2006). By contrast, the male germ line seems to normally exclude the mitochondrial F mitochondrial genome and preferentially amplify the M genome (Venetis *et al.*, 2006). Indeed, by forcing spawned sperm to swim through a solution of Percoll, and thus minimizing the probability of sperm contamination by somatic tissues or cells, Venetis *et al.* (2006) recovered only distinct, paternally transmitted mtDNA genomes in the total DNA extractions from the 'washed' sperm of 36 *M. galloprovincialis* individuals. This precise male-specific transmission of the M mitochondrial genome is required for the stability of DUI (Venetis *et al.*, 2006).

Another unusual aspect of mussel genetics (and to date this has only been demonstrated in *Mytilus*) is that some females produce female-biased offspring whereas other females produce male-biased offspring, regardless of which male they mate with (Saavedra, Reyero and Zouros, 1997). A model to explain the coupling of mtDNA inheritance and sex determination in mussels has been previously developed (Saavedra, Reyero and Zouros, 1997; Zouros, 2000; Kenchington *et al.*, 2002; Cogswell, Kenchington and Zouros, 2006) (Box 1). According to this model, mtDNA inheritance in bivalves with DUI could be controlled by a maternally-encoded sex-determining gene or a gene linked to a sex-determining factor, as has been observed in the basidiomycete fungus *Cryptococcus neoformans* (Xu, 2005). It is tempting to speculate that the sperm mitochondria might be involved in differentiation of male reproductive tissue once they have been sequestered in the germ line. This hypothesis would imply functional differences between the mitochondrial M and F genomes and/or specific nucleo-mitochondrial interactions in the male gonad, although to our knowledge such interactions have not yet been demonstrated.

Box 1. DUI model

The model, based on sex-ratio bias of certain pair matings, proposes that the bias is under the control of the female parent and suggests that this control is exercised by the nuclear genotype of the mother rather than mitochondrial genotype (Kenchington *et al.*, 2002). It has been hypothesized that three genetic factors (i.e. W, X, and Z) are involved. According to this model, W is located in the outer surface of the sperm mitochondrion and is recognized by an egg nuclear-encoded cytoplasmic X factor. These two

factors are suggested to be parts of the 'sperm mitochondria elimination system' that leads to maternal mtDNA inheritance, a mechanism that could be comparable to the ubiquitination system observed in mammals (Kenchington *et al.*, 2002). In addition to W and X, the model also implies a DUI-specific, one-locus two-allele, Z factor (i.e. the active Z and the inactive z), which occurs in the egg cytoplasm. The role of the factor Z is to suppress factor X, and thus prevent the elimination (or dispersion) of sperm mitochondria in fertilized eggs (Figure 1). If only sperm mitochondria subsequently gain entrance into the primordial germ cell, or if the M genome has a replicative advantage over the F genome in these cells, it might explain the dominance of the M genome in the male gonad (Cogswell, Kenchington and Zouros, 2006). The Z factor, paternal mitochondria and the M genome herein are virtually linked to sex determination.



Box 1 Figure I. Proposed genetic control of mitochondrial transmission under the DUI system. (a) Male and female gametes contribute to the mitochondrial population of the fertilized egg (b) Mother-dependant genetic models (Saavedra, Reyero and Zouros, 1997; Zouros, 2000; Dalziel and Stewart, 2002; Cogswell, Kenchington and Zouros, 2006) predict that mothers carrying the Z allele will produce eggs with the Z factor, allowing the retention of sperm mitochondria and their subsequent aggregation in a single blastomere in embryos destined to become males. By contrast, zz mothers will produce eggs without the Z factor in which sperm mitochondria will be dispersed and/or lost, and embryos will become females.

2.4 Molecular evolution of M and F mitochondrial genomes

Genetic analyses using partial mtDNA sequences of M and F mitotypes in species with DUI (Rawson and Hilbish, 1995; Stewart *et al.*, 1995; Hoeh *et al.*, 1996b; Liu, Mitton and Wu, 1996; Curole and Kocher, 2002; Passamonti, Boore and Scali, 2003; Curole and Kocher, 2005) indicated that (i) DUI appeared 200 million years ago, if not before; (ii) mussel mtDNA (both M and F lineages) has experienced an accelerated rate of mtDNA sequence divergence compared with that of other animal taxa; and (iii) the M genome evolves more quickly than the F genome. A leading interpretation explaining why M evolves faster than F, and why both of them evolve faster than typical mtDNA, is the relaxation of selective constraints due to the unequal 'division of labour' in the DUI system (Stewart *et al.*, 1996; Zouros, 2000). Indeed, contrary to typical mtDNA, which has to perform fully in gonad and soma of both sexes, the F genome functions in female gonad and soma of both sexes, whereas the male genome serves primarily in the male gonad and only partially in male soma, where it occurs sporadically in conjunction with the more abundant F type (Stewart *et al.*, 1996; Garrido-Ramos *et al.*, 1998; Dalziel and Stewart, 2002).

The availability of 11 complete or nearly complete F and M mtDNA genome sequences has shown how species with DUI vary from the typical pattern of animal mtDNA gene content (Table 1). Most species with DUI [except *Lampsilis* (a unionoidean bivalve)] have lost the gene for ATPase subunit 8 (*atp8*) and some have a second tRNA gene for methionine (*trnM*) (e.g. *Mytilus* spp. and female *Venerupis*). However, other bivalve species, such as the sea scallop, *Placopecten magellanicus*, also lack *atp8* and have multiple *trnM* genes (Smith and Snyder, 2007).

Table 2.1 Complete mitochondrial genomes of species with DUI.

Species	Gender	Order	Genome Size (bp)	Gene Number (proteins-tRNAs- rRNAs)	GenBank accession number	Reference
Marine mussels						
<i>Mytilus edulis</i>	Female	Mytiloida	16,740	12 – 23 – 2	AY484747	(Boore, Medina and Rosenberg, 2004); (Hoffmann, Boore and Brown, 1992)
<i>M. galloprovincialis</i>	Female	Mytiloida	16,744	12 – 23 – 2	AY497292	(Mizi et al., 2005)
<i>M. edulis</i>	Male Haplotype 1	Mytiloida	16,622	12 – 23 – 2	AY823623	(Breton et al., 2006)
<i>M. edulis</i>	Male Haplotype 2	Mytiloida	16,624	12 – 23 – 2	AY823624	(Breton et al., 2006)
<i>M. galloprovincialis</i>	Male	Mytiloida	17,671	12 – 23 – 2	AY363687	(Mizi et al., 2005)
<i>M. trossulus</i>	Masculinized	Mytiloida	18,652	12 – 23 – 2	AY823625	(Breton et al., 2006)
Marine clam						
<i>Venerupis (Ruditapes) philippinarum</i>	Female	Veneroida	22,676	13 – 23 – 2 ^b	AB065375	M. Okazaki and R. Ueshima,

<i>V. philippinarum</i>	Male	Veneroida	21,441	12 – 24 – 2	AB065374	unpublished M. Okazaki and R. Ueshima, unpublished
Freshwater mussels						
<i>Lampsilis ornata</i>	Female	Unionoida	16,060	13 – 22 – 2 ^c	AY365193	(Serb and Lydeard, 2003)
<i>Inversidens japanensis</i>	Female	Unionoida	16,826 ^a	12 – 22 – 2	AB055625	M. Okazaki and R. Ueshima, unpublished
<i>I. japonensis</i>	Male	Unionoida	16,967 ^a	12 – 22 – 2	AB055624	M. Okazaki and R. Ueshima, unpublished

^aIncomplete mitochondrial sequences. ^bCox2 is duplicated. ^cContains *atp8*.

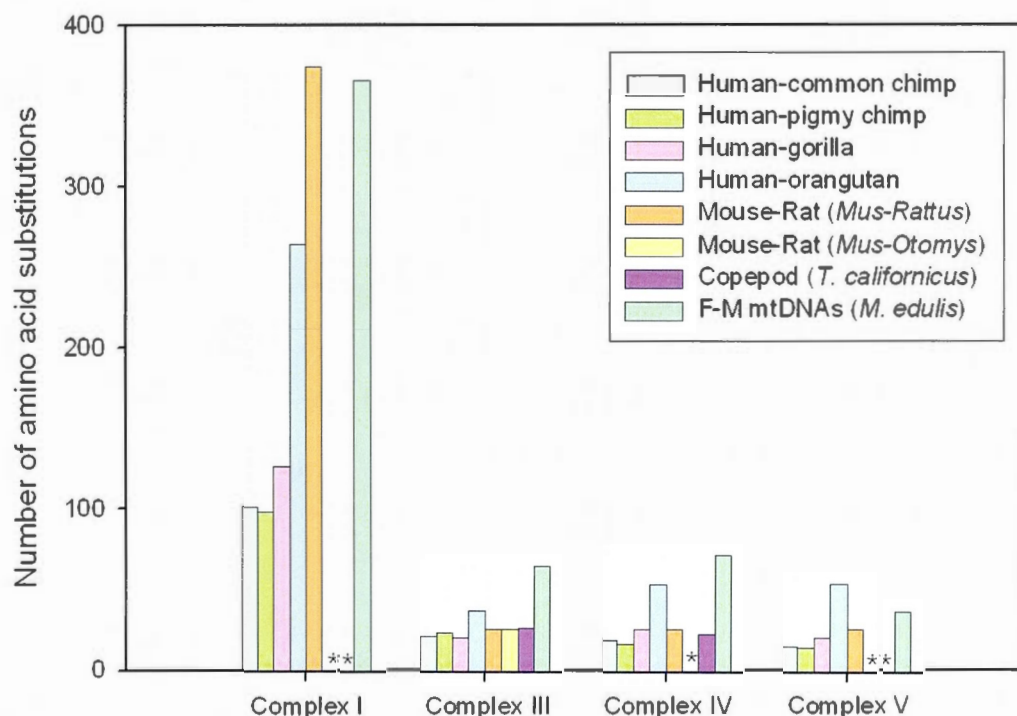
The mtDNA control region (i.e. the region of the mitochondrial genome in which replication and transcription are initiated) has been identified in mytilid bivalves (Cao *et al.*, 2004), but, to our knowledge, no control region has been confirmed yet in veneroid or unionoidean bivalves [see Ref. (Serb and Lydeard, 2003)]. With few exceptions, gene order and content of F and M genomes from the same species are well preserved. The freshwater mussel *Inversidens japonensis* has two gene-order inversions (in both the light and the heavy strands) and two tRNAs (trnD and trnV) encoded by opposite strands that are responsible for differences between F and M mtDNAs (Mizi *et al.*, 2005) (M. Okazaki and R. Ueshima, unpublished). The F genome of the marine clam *Venerupis philippinarum* contains a duplication of the gene for cytochrome c oxidase subunit II (cox2) and the M genome contains an extra trnM gene, distinguishing the two mtDNAs (Mizi *et al.*, 2005) (M. Okazaki and R. Ueshima, unpublished).

A recent comparative analysis of several mitochondrial genomes from three *Mytilus* species showed that the amino acid variability within mtDNA regions of the M genome was highly correlated with variability within the F genome (i.e. regions with high or low amino acid differences were similar in both lineages). This was interpreted as evidence of cyto-nuclear co-evolution (Box 2). The necessity of evolving in the same nuclear background has apparently forced the F and M mtDNA genomes to experience similar selective pressures. These selective pressures could result in the retention of a particular amino acid from the ancestral mtDNA genome at positions of structural importance to enzyme function, convergent amino acid changes at other sites, or even functionally equivalent (but different) amino acid substitutions in the M and F lineages. The key observation from this study is

that the regions of variability in the distinct gender-associated mtDNA lineages are highly correlated (Breton *et al.*, 2006).

Box 2. DUI and the study of intergenomic interactions

Strong evidence for intergenomic co-evolution has been provided by direct manipulation of cells and/or embryos in culture or laboratory crosses where OXPHOS functional assays can be linked to amino acid substitutions in the interacting proteins encoded by mitochondrial and nuclear genes [see Refs Blier, Dufresne and Burton (2001) and Rand, Haney and Fry (2004) for reviews]. Levels of intra-species mtDNA sequence divergences between M and F genomes in marine and freshwater male mussels are the highest intra-individual values yet reported (with sequence divergences in *cox1* often >20% and >30%, respectively), even higher than intra- or inter-species values reported in classical model systems used for the study of intergenomic co-evolution (Figure 1) [see Ref. Blier *et al.*, 2006]. From a co-evolutionary perspective, one challenging question is: 'Do nuclear-encoded peptides function equally-well with either the M or F mtDNA?'. For example, there is the possibility that in somatic tissue of either sex the M genome is present but not serving any useful function. Another hypothesis is that the M genome might do its job even under a clumsy collaboration with the nuclear genes. A third assumption is the existence of nuclear isoforms that are only expressed and interact with either the M or F mtDNA genome. These isoforms might exhibit differences in biochemical activity that could have evolved to avoid potential within-individual intergenomic conflict. Indeed, testis-specific isoforms (e.g. nuclear-encoded mitochondrial cytochrome c and COX subunit VIb [complex IV]) have been discovered in mammals, suggesting a possible accommodation for the high energy demand of sperm motility (Das, 2006).



Box 2. Figure 1. Divergences in model systems for the study of co-evolution. Amino acid sequence differences of mtDNA-encoded peptides in model systems for analyses of cytonuclear co-evolution in which disruption of respiratory chain function has been observed. Human cells with mtDNA from chimpanzee and gorilla had impaired oxygen consumption (i.e. decreased, on average, by 20%, 34%, and 27%, respectively, compared with the human cells and/or human mtDNA control) that was attributable to a marked deficiency in respiratory chain complex I (Kenyon and Moraes, 1997; Barrientos, Kenyon and Moraes, 1998). Activities of complexes I, III, and IV in the *Rattus* xenocybrid were 44%, 37%, and 78% of control mean, respectively, whereas the activity of complex III was 2% of control cybrid in the *Otomys* xenocybrid (McKenzie and Trounce, 2000; McKenzie *et al.*, 2003). COX activities were decreased as mitochondria from a population were moved to a purer and purer nuclear background from a distant

population (Burton, Rawson and Edmands, 1999; Ellison and Burton, 2006).*, Data not available.

Because of the important differences in the mutation rates of male and female mtDNAs, it has been hypothesized that male mtDNA should be either under relaxed constraints or positive selection pressures. In both cases, these differences can offer exceptional material to reveal, by comparative approaches (i.e. male and female mtDNAs of different species), the hot-spots of selective pressures under common nuclear backgrounds. Such studies could be of major significance because most models of mtDNA evolution consider neutral processes and purifying selection as the major forces shaping the pattern of sequence divergence among species.

2.5 When DUI breaks down: masculinization of F mtDNA

Phylogenetic analyses of cytochrome c oxidase subunit I (cox1) sequences have demonstrated that in marine (but not freshwater) mussels, the fidelity of DUI is sometimes compromised. Some males seem to lack a typical M genome (Hoeh *et al.*, 1997; Quesada, Wenne and Skibinski, 1999; Zouros, 2000; Quesada, Stuckas and Skibinski, 2003), and F genomes seem to occasionally invade the male route of inheritance such that they become transmitted from generation to generation only through sperm (Saavedra, Reyero and Zouros, 1997; Zouros, 2000; Obata and Komaru, 2005; Venetis *et al.*, 2006). However, the reverse (i.e. an M genome invading the F genome route of transmission) has not been observed. This phenomenon has been referred to as a 'masculinization' or 'role-reversal' event. The F genomes that have invaded the M genome's route of transmission are referred to as 'recently-masculinized' M types (Hoeh *et al.*, 1997; Quesada, Wenne and Skibinski, 1999; Zouros, 2000; Quesada, Stuckas and Skibinski, 2003).

Several populations of *Mytilus* mussels are polymorphic for two classes of M mitochondrial genomes: an older, 'standard' M type and a 'recently-masculinized' M type. The genomes of the latter, particularly their protein-coding regions, are highly similar to F genomes, but they are transmitted as M genomes through sperm.

There are at least two potential explanations for the breakdown of DUI in *Mytilus* spp. According to one view, failure of DUI is associated with hybridization events between pairs of *Mytilus* species (Hoeh, Stewart and Guttman, 2002; Wood *et al.*, 2003; Curole and Kocher, 2005), leading to the disruption of nuclear–cytoplasmic interactions and DUI instability (Curole and Kocher, 2005). A second explanation for the existence of this masculinization phenomenon came with the discovery of intermolecular mitochondrial recombination within the *cox3* gene in male gonadal tissues from *Mytilus galloprovincialis* (Ladoukakis and Zouros, 2001). It should be noted that both hypotheses are not mutually exclusive; recombination might occur more frequently (or perhaps is more easily detected) in hybrid situations. These findings are highly significant because they provided direct evidence of recombination in animal mtDNA, a much debated subject [reviewed in Ref. (Rokas, Ladoukakis and Zouros, 2003)]. Mitochondrial recombination was later confirmed in other mytilid and venerid species (Burzynski *et al.*, 2003; Passamonti, Boore and Scali, 2003; Rawson, 2005; Breton *et al.*, 2006; Burzynski *et al.*, 2006). The observation that some PCR-amplified main control regions in male gonads of *M. trossulus* were a mosaic of F- and M-like sequences provided a potential link between the homologous recombination of F and M genomes and masculinization (Burzynski *et al.*, 2003). Specifically, because these recombinant variants were transmitted through sperm like the M genome, but showed high coding sequence similarity to the

M. edulis F genome, Burzynski *et al.* (2003) speculated that occasional invasions of the male transmission route by the F genome could be possible through the addition of M control region sequences to the control region of F genomes. The first complete sequence of a recently-masculinized *M. trossulus* mitochondrial genome also indicated that two control regions exist, one male-specific and the other female-specific (Breton *et al.*, 2006). At present, it is not known whether both potential control regions are functional, whether they act in a tissue-specific manner, or what their respective roles are in mitotype transmission. Nevertheless, recent evidence of mitochondrial recombination within the control region in male and female *M. trossulus* mussels corroborates the hypothesis that an M-type control region sequence is necessary to confer the paternal role on genomes that are otherwise F-like (Burzynski *et al.*, 2006).

Interestingly, there is no evidence of mitochondrial recombination or masculinization events during the evolutionary history of freshwater mussels (Unionoida) (Hoeh, Stewart and Guttman, 2002; Walker *et al.*, 2006). The absence of recombination could explain why gender-switching events are lacking in unionoidean bivalves (Mizi *et al.*, 2005). The lack of masculinization in freshwater mussels also coincides with the presence of a unique M genome-specific 3' extension of the mtDNA-encoded cytochrome c oxidase subunit II gene (Mcox2) (Curole and Kocher, 2002; Curole and Kocher, 2005; Chakrabarti *et al.*, 2006). This extension, which is 185 codons in length, is present in all unionoidean M genomes cox2 genes examined to date, including each of the three unionoidean bivalve families (i.e. Hyriidae, Margaritiferidae and Unionidae) (Curole and Kocher, 2002; Curole and Kocher, 2005). Examination of the rates and patterns of substitution suggests that the extension (Mcox2e) is evolving under relaxed purging selection

relative to the upstream Mcox2 homologous region (Mcox2h), that is, the region present in both Fcox2 and Mcox2 (Curole and Kocher, 2002; Curole and Kocher, 2005). The Mcox2e is likely to be the most rapidly evolving mitochondrial domain identified in animals (Curole and Kocher, 2005). Apparently, Mcox2e is neither present in mytiloid M genomes, nor in other animal mitochondrial genomes (Mizi *et al.*, 2005; Breton *et al.*, 2006).

A specific function for Mcox2e has not yet been demonstrated. The cox2 gene encodes a highly conserved subunit of cytochrome oxidase, the terminal enzyme of the mitochondrial inner membrane that is responsible for the transfer of electrons from cytochrome c to oxygen (Rawson and Burton, 2006). MCOX2e antibody-based analyses indicated that the extended MCOX2 protein in *Venustaconcha ellipsiformis* (Unionoidea:Unionidae) is predominantly expressed in testes, weakly expressed in other male tissues, and not expressed in female tissues (Chakrabarti *et al.*, 2006). The immunohistochemistry-based localization of MCOX2 to sperm mitochondria combined with the predicted presence of five transmembrane helices in the *V. ellipsiformis* MCOX2e region suggest that the latter is located in the outer and/or inner mitochondrial membrane (Chakrabarti *et al.*, 2006). These characteristics suggest several potential functions for the MCOX2e; for example, within the sperm mitochondria and in developmental interactions. During spermatogenesis, Mcox2e expression could be involved in apoptosis, which is an important physiological mechanism that regulates the number of sperm produced [e.g. Chan *et al.*, 2006]. One likely possibility for a developmental interaction functionality for MCOX2e is that an outer mitochondrial membrane localization could facilitate gender-specific movements of sperm-derived mitochondria within unionoidean embryos in a manner similar to that observed in *Mytilus* (Cao, Kenchington and Zouros,

2004; Cogswell, Kenchington and Zouros, 2006). Although more studies are needed to elucidate the function(s) of the MCOX2e protein, its association with the absence of masculinization in unionoidean bivalves suggests that it has been selected either as a protective mechanism against gender-switching or for advantageous male reproductive function (which could also explain the lack of gender-switching).

2.6 Origin and evolution of DUI

Even if the occurrence of DUI in other taxa remains to be explored, the question persists why DUI evolved in bivalve mollusks (Burt and Trivers, 2006; Davison, 2006). Can the considerable variation of bivalve reproductive strategies provide a clue (Paz *et al.*, 2005)? Although individuals of many bivalve species demonstrate stable gonochoric sexuality (i.e. once they become male or female, they remain that sex throughout their lives), several species have simultaneous hermaphrodites that produce both male and female gametes in the ovotestis or sequential hermaphrodites that change sex as they age. In all cases, sex (or type of gamete produced) is not determined until germ cells are differentiated but the exact mechanism(s) of sexualization of the undifferentiated gonad are unknown (Paz, Mikhailov and Torrado, 2001). Moreover, neither sex chromosomes nor sex-related genes have been identified, except for the detection of an esterase (Est)-like 'male-associated polypeptide' in the male gonad (or male reproductive tissue), and a fibronectin (Fn)-like polypeptide in the female gonad (or female reproductive tissue) of the gonochoric DUI-species *Mytilus galloprovincialis* and the hermaphroditic *Pecten maximus* (Torrado *et al.*, 2003; Paz *et al.*, 2005).

A recent hypothesis suggests that DUI first emerged in a simultaneously hermaphroditic species in which both eggs and sperm were produced in an

ovotestis (Davison, 2006). This idea is based on the assumption that in species with (i) distinct sex chromosomes, (ii) maternal transmission of mtDNA, and (iii) heterogametic males, maternally sex-linked genes and maternally inherited mitochondria should be co-adapted (i.e. mitochondria are more frequently co-transmitted with maternally expressed X-linked genes than with autosomal genes, whereas they are never co-transmitted with paternally expressed Y-linked genes) (Rand, 2001). In this context, it is not in the interest of maternally expressed sex-linked genes to 'allow' paternal inheritance to persist (Davison, 2006). By contrast, no such conflict exists in hermaphrodite species, which do not have sex chromosomes and in which both male and female gametes are produced in the ovotestis. Perhaps paternally inherited mitochondria were retained for some time in a hermaphroditic ancestral species; however, to persist over the longer term, paternal M genomes would have been selected to secure more reliable transmission. This could have been achieved by the masculinization of the ovotestis (i.e. suppression of its ovarian aspects), so that paternal transmission of the M genome and a dioecious sex-determination system would become associated (see above) (Davison, 2006). However, further studies on bivalve species with and without DUI, as well as on hermaphroditic and dioecious species will be essential to better understand the coupling of gender and mtDNA inheritance in the DUI system.

2.7 Potential adaptive evolution of M mtDNA

Currently, it is unclear whether a selective advantage favoured the retention of DUI in bivalves, but one possible mechanism could be related to the sex-antagonistic effects of mtDNA resulting from maternal inheritance (Rand, 2001; Ballard and Whitlock, 2004; Burt and Trivers, 2006). In SMI,

natural selection on mtDNA operates only in females because males do not transmit their cytoplasmic genes (Zeh, 2004; Zeh and Zeh, 2005). Consequently, mtDNA genotypes that have positive (or neutral, or even slightly deleterious) fitness effects in females but potentially deleterious effects in males, can theoretically be maintained in a population (Frank and Hurst, 1996; Rand, 2001; Gemmell, Metcalf and Allendorf, 2004). For example, it has been shown that the human mtDNA variant T, which is observed at a frequency of 20% in some European populations, yields significantly less motile sperm than the most common and best performing human mtDNA variant H (Ruiz-Pesini *et al.*, 2000). Recent experimental studies of mice provide even more compelling evidence that mitochondrial mutations with comparatively small effects on female function can have profound impacts on spermatogenesis (Trifunovic *et al.*, 2004) and the incidence of major sperm abnormalities (Nakada *et al.*, 2006). These findings provide strong support for the hypothesis that adaptive evolution of male function might be significantly constrained because of the maternal inheritance of mitochondria (Frank and Hurst, 1996). Doubly uniparental inheritance of mitochondria would thus be both an elegant strategy to avoid sex-specific constraints associated with maternal mtDNA transmission and an opportunity for mitochondria to evolve adaptively for male function (Gemmell and Braisher, 2001; Selosse, Albert and Godelle, 2001; Zeh and Zeh, 2005).

An interesting approach to address the issue of adaptive evolution of M genome in males employs the comparison of functional properties of mitochondria between 'standard' and 'recently-masculinized' mytilid mussel spermatozoa. Because these different male mtDNAs exhibits almost 9% amino acid sequence divergence, and that many of these amino acid substitutions are not conservative, it was hypothesized that these differences

could affect mitochondrial functions, and thus sperm motility (Everett *et al.*, 2004). The initial test of this hypothesis indicated no significant differences in sperm swimming speeds between standard male and recently masculinized mitotypes in *Mytilus edulis* (Everett *et al.*, 2004). Swimming speed is, however, only one of the key parameters of sperm motility and fitness, and we can not exclude the possibility of more subtle effects on mitochondrial respiratory chain function or on other gamete characteristics (e.g. longevity, numbers of sperm). Additional comparative analyses of masculinized and standard M-type bearing-sperm are needed to clarify the potential impact associated with the amino acid substitutions observed between these mtDNAs.

2.8 Concluding remarks

Since its discovery in the early 1990s, much progress has been made in understanding the DUI system in bivalve species, in particular with regard to the mechanisms underlying the sex-specific behavior of sperm mitochondria and the molecular evolution of M and F mitochondrial genomes. Future work will focus on unique features of the DUI system such as the potential adaptive evolution of the M genome on sperm motility, the role of recombination in masculinization events (and specifically, the role of the control region in determining whether a particular mitochondrial genome will behave as an F or an M genome), functional aspects of the exclusive Mcox2 extension, and finally, the possible association between mtDNA inheritance and sex determination. As these examples illustrate, DUI provides an excellent opportunity to investigate evolutionary and functional consequences of alternative modes of mitochondrial inheritance.

2.9 Acknowledgements

We thank Gertraud Burger and the anonymous reviewers for suggestions that improved this article. This work was supported by research grants from the National Sciences and Engineering Research Council (NSERC) to P.U. Blier and D.T. Stewart, and from the National Science Foundation (NSF) to W.R. Hoeh. S. Breton and H. Doucet Beaupré were financially supported by the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) and NSERC scholarships.

2.10 Glossary

- Gonochoric describes a sexually reproducing species in which there are two (at least) distinct sexes.
- Heteroplasmy: the existence of two (or more) plastid variants (mitochondrial or chloroplast DNA) within an organelle, cell, tissue or individual.
- Homoplasmy: the condition in which all plastid genomes (i.e. usually referring to genetic identity of mitochondria or chloroplasts) in an organelle, cell, tissue or individual are identical.
- Hybrid zone: an area where two species come into contact and offspring are produced that are the result of interbreeding between the different species.
- Masculinization: a female-transmitted mtDNA can become 'masculinized' (i.e. reverse its role and become transmitted paternally).
- Nuclear isoforms: different forms of a nuclear-encoded protein that might be produced from different genes or from the same gene by

alternative splicing (i.e. different versions of messenger RNA) or posttranslational modifications. Isoforms are usually tissue-specific.

- Ovotestis refers to a gonad that contains both ovarian and testicular tissues (i.e. hermaphroditism).
- Oxidative phosphorylation (OXPHOS): the synthesis of ATP (i.e. the high energy source used for essentially all active metabolic processes within the cell) by phosphorylation of ADP for which energy is obtained by electron transport and which takes place in the mitochondria during aerobic respiration.
- Respiratory chain complexes and ATP synthase: an elaborate system composed of five enzyme complexes situated on the inner mitochondrial membrane. The four first complexes act as an electron transport chain from reduced cofactors (e.g. the reduced form of nicotinamide adenine dinucleotide [NADH] or flavin adenine dinucleotide [FADH₂]) to molecular oxygen. The passage of electrons is linked to proton efflux across the mitochondrial inner membrane to establish a source of power for ATP synthase (complex V) to synthesize ATP. The mtDNA encodes seven subunits of the NADH: ubiquinone oxidoreductase (complex I; ND1 to ND6, and ND4L), one subunit of the ubiquinone: cytochrome c oxidoreductase (complex III; Cyt b), three subunits of the cytochrome c oxidase (complex IV; COX1 to COX3) and two subunits of the ATP synthase (complex V; ATP6 and ATP8). Other subunits of complexes I, III, IV, and V, as well as all the components of complex II (succinate: ubiquinone oxidoreductase), the membrane transporters, the enzymes of the matrix, and the factors involved in other mitochondrial functions, e.g. mtDNA replication and mtDNA expression, are nuclear-encoded in animals.

- Selfish mitochondrial elements: potentially deleterious elements that enhance their own transmission relative to their allelic counterparts.

CHAPITRE III

MITOCHONDRIAL PHYLOGENOMICS OF THE BIVALVIA (MOLLUSCA): SEARCHING FOR THE ORIGIN AND MITOGENOMIC CORRELATES OF DOUBLY UNIPARENTAL INHERITANCE OF mtDNA

Hélène Doucet-Beaupré¹, Sophie Breton², Eric G. Chapman³, Pierre U. Blier¹, Arthur E. Bogan⁴, Donald T. Stewart⁵ and Walter R. Hoeh²

¹Département de Biologie, Université du Québec à Rimouski, 300 Allée des Ursulines, Rimouski, Québec, G5L 3A1, Canada (H. Doucet-Beaupré : conception du projet, expérimentation et analyse des données, rédaction de l'article et conception des figures ; P. U. Blier : supervision et aide à la rédaction).

²Department of Biological Sciences, Kent State University, Kent, Ohio 44242, USA (S. Breton: expérimentation et analyse des données, aide à la rédaction ; W. R. Hoeh : aide à la rédaction)

³Department of Entomology, University of Kentucky, Lexington, Kentucky, 40546-0091 USA (E.G. Chapman: aide au traitement des données et à la rédaction)

⁴North Carolina State Museum of Natural Sciences, Research Laboratory, MSC 1626, Raleigh, North Carolina, 27699-1626 USA (A. E. Bogan: aide à la collecte de données)

⁵Department of Biology, Acadia University, 33 Westwood Ave, Wolfville, NS, B4P 2R6, Canada (D.T. Stewart: aide à la rédaction)

Paru en 2010 dans BMC Evolutionary Biology 10:50

Résumé

Contexte : La double hérédité uniparentale (DUI) est un système d'hérédité mitochondriale atypique retrouvé uniquement chez certains bivalves. À l'intérieur de la DUI, l'ADNmt maternel (génom F) et paternel (génom M) donne deux lignées mitochondriales distinctes associées au sexe. C'est dans les moules d'eau douce (de l'ordre des Unionoida) que l'on rencontre les lignées génomiques F et M les plus anciennes. Des analyses comparatives des génomes mitochondriaux des Unionoïdés combinées à un solide cadre phylogénétique sont nécessaires pour élucider l'origine, la fonction et l'évolution moléculaire de la DUI. Dans la présente étude, les génomes F et M de trois espèces d'Unionoïdés, *Venustaconcha ellipsiformis*, *Pyganodon grandis* et *Quadrula quadrula* ont été séquencés. Des analyses génomiques comparatives ont été réalisées sur ces six génomes avec deux génomes d'Unionoïdés F et un génome M provenant de GenBank (les génomes F et M d'*Inversidens japonensis* et le génome F de *Lampsilis ornata*).

Résultats : Comparativement à leurs homologues F, les génomes M d'Unionoïdés contiennent des caractéristiques uniques, y compris une localisation nouvelle du gène *tmH*, une inversion des gènes *atp8-tmD* et une extension codante unique du gène du cytochrome c oxydase sous-unité II en 3'. Une ou plusieurs de ces caractéristiques uniques du génome M pourraient avoir un lien de causalité associé à la transmission paternelle. Les bivalves unionoïdes sont caractérisés par une extrême divergence intraspécifique entre les séquences d'ADNmt associé au sexe avec une moyenne de 50% pour *V. ellipsiformis*, 50% pour *I. japonensis*, 51% pour *P. grandis* et 52% pour *Q. quadrula* (distances *p* en acides aminés non corrigées). Les analyses phylogénétiques de 12 gènes codant de 29 bivalves et cinq génomes mitochondriaux comme groupe externe indiquent une monophylie robuste des bivalves et suggèrent l'ordre suivant de ramification dans les bivalves autolamellibranches : [[Pteriomorphia + Veneroida] + Unionoida].

Conclusion : La nature basale des Unionoïdés au sein des bivalves autolamellibranches et l'hypothèse de l'origine unique de la DUI suggèrent que (1) la DUI est apparue dans une lignée ancestrale des bivalves autolamellibranches et aurait ensuite été perdu dans de multiples lignées subséquentes (2) les caractéristiques des génomes mitochondriaux observées chez les bivalves unionoïdes pourraient ressembler davantage à la condition ancestrale de la DUI. Les descriptions et les comparaisons présentées dans cet article sont essentielles à une compréhension plus complète de l'origine et des conséquences de la DUI.

3.1 Abstract

Background

Doubly uniparental inheritance (DUI) is an atypical system of animal mtDNA inheritance found only in some bivalves. Under DUI, maternally (F genome) and paternally (M genome) transmitted mtDNAs yield two distinct gender-associated mtDNA lineages. The oldest distinct M and F genomes are found in freshwater mussels (order Unionoida). Comparative analyses of unionoid mitochondrial genomes and a robust phylogenetic framework are necessary to elucidate the origin, function and molecular evolutionary consequences of DUI. Herein, F and M genomes from three unionoid species, *Venustaconcha ellipsiformis*, *Pyganodon grandis* and *Quadrula quadrula* have been sequenced. Comparative genomic analyses were carried out on these six genomes along with two F and one M unionoid genomes from GenBank (F and M genomes of *Inversidens japonensis* and F genome of *Lampsilis ornata*).

Results

Compared to their unionoid F counterparts, the M genomes contain some unique features, including a novel localization of the trnH gene, an inversion of the atp8-trnD genes and a unique 3'coding extension of the cytochrome c oxidase subunit II gene. One or more of these unique M genome features could be causally associated with paternal transmission. Unionoid bivalves are characterized by extreme intraspecific sequence divergences between gender-associated mtDNAs with an average of 50% for *V. ellipsiformis*, 50% for *I. japonensis*, 51% for *P. grandis* and 52% for *Q. quadrula* (uncorrected amino acid p-distances). Phylogenetic analyses of 12 protein-coding genes from 29 bivalve and five outgroup mt genomes robustly indicate bivalve monophyly and the following branching order within the autolamellibranch bivalves: [[Pteriomorpha + Veneroida] + Unionoida].

Conclusion

The basal nature of the Unionoida within the autolamellibranch bivalves and the previously hypothesized single origin of DUI suggest that (1) DUI arose in the ancestral autolamellibranch bivalve lineage and was subsequently lost in multiple descendant lineages and (2) the mitochondrial genome characteristics observed in unionoid bivalves could more closely resemble the DUI ancestral condition. Descriptions and comparisons presented in this paper are fundamental to a more complete understanding regarding the origins and consequences of DUI.

3.2 Background

Mitochondrial DNA (mtDNA) is the only extranuclear genome in the animal cytoplasm. Located in the matrix of mitochondria, metazoan mtDNA is normally a small circular DNA molecule about 14-16 kilobases (kb) long usually encoding the same 37 genes [Attardi (1985); Boore (1999); but see Gissi, Iannelli and Pesole (2008) for exceptions]. Typically, all mtDNAs in the zygote come from the oocyte and even though evidence for occasional paternal leakage has been reported (Xu, 2005; Elson and Lightowers, 2006), animal mtDNA is thought to strictly follow maternal inheritance (Birky, 2001). This clonal inheritance coupled with the successive cell divisions that represent sequential bottlenecks for the mitochondrial population (Birky, 2001; Rand, 2001; Shoubridge and Wai, 2008) result in an essentially homoplasmic state for mtDNA. An extreme exception to the paradigm of strict maternal inheritance of animal mtDNA (SMI) is found in three bivalve lineages (i.e., the subclasses Pteriomorpha: Order Mytiloida; Paleoheterodonta: Order Unionida and Heterodonta: Order Veneroida), which possess an unusual system termed doubly uniparental inheritance of mtDNA (DUI) [see Breton *et al.* (2007) and Passamonti and Ghiselli, (2009) for reviews].

In DUI-possessing organisms, distinct gender-associated mitochondrial DNA lineages coexist: a female-transmitted (F) genome and a male-transmitted (M) genome. Under DUI, female bivalves transmit their mitochondria (carrying F mtDNA) to both sons and daughters, as in SMI, but males pass on their mitochondria (via sperm carrying M mtDNA) to only sons [e.g., Sutherland *et al.* (1998) but see Obata *et al.*, (2007)]. At the organismal level, male bivalves with DUI are thus heteroplasmic and contain both M and F genomes. In male somatic tissues, the F genome predominates while in

male gonadal tissues, the M genome is predominant (Garrido-Ramos *et al.*, 1998) and it appears to be the exclusive type in sperm (Venetis *et al.*, 2006). In females, both somatic and gonadal tissues typically contain the F genome, but the occasional presence of a small amount of the M genome has been demonstrated in somatic tissues and ovaries of some species (Garrido-Ramos *et al.*, 1998; Obata *et al.*, 2006; Venetis *et al.*, 2006; Chakrabarti *et al.*, 2007; Obata *et al.*, 2007).

The broad taxonomic distribution of DUI within the Bivalvia [e.g., Stewart *et al.*, 1995; Hoeh *et al.*, 1996b; Passamonti and Scali, 2001; Curole and Kocher, 2002; Hoeh, Stewart and Guttman, 2002; Passamonti, Boore and Scali, 2003; Walker *et al.*, 2006; Theologidis *et al.*, 2008] supports the idea that it evolved once in an ancestral bivalve lineage, from standard uniparental inheritance, and was lost in some descendant bivalve lineages (e.g., oysters and probably scallops) (Zouros, 2000; Hoeh, Stewart and Guttman, 2002; Obata *et al.*, 2008). DUI could then be the ancestral condition for the Bivalvia. However, a more definitive statement to this effect rests on producing a more reliable bivalve phylogeny along with clarifying the distribution of DUI in additional bivalve lineages. Although many of the essential elements of DUI have been described, [i.e., distinct M and F lineages, heteroplasmy in males, rapid molecular evolution, particularly of M types (Rawson and Hilbish, 1995; Stewart *et al.*, 1995; Hoeh *et al.*, 1996a; 1996b; Liu, Mitton and Wu, 1996; Stewart *et al.*, 1996; Passamonti and Scali, 2001; Curole and Kocher, 2002; Passamonti, Boore and Scali, 2003; Breton *et al.*, 2006], the current and/or historical function of DUI still remains unknown. Comparisons of entire F and M genomes (in contrast to partial sequences of a few genes) will enable the characterization of potential gene content/organizational/functional differences between the M and F genomes, and will help to reconstruct the

history of any possible recombination and/or gene translocation events in these distinct, gender-associated lineages.

To date, 15 complete or nearly complete F and M mtDNA genome sequences are available for species with DUI but these are numerically biased towards marine taxa (i.e., species from the Mytiloida and Veneroida) (Hoffmann, Boore and Brown, 1992; Serb and Lydeard, 2003; Boore, Medina and Rosenberg, 2004; Mizi *et al.*, 2005; Venetis *et al.*, 2007; Zbawicka, Burzynski and Wenne, 2007) (Table 2.1). While the vertebrate mitochondrial gene order is almost invariant, mollusks, and bivalves in particular, exhibit radical rearrangements of mitochondrial genes and extensive mtDNA variability at the intrageneric level (Attardi, 1985; Vallès and Boore, 2006; Smith and Snyder, 2007). For example, the two congeneric oyster species *Crassostrea virginica* and *C. gigas*, both lacking DUI (Obata *et al.*, 2008), show broad differences in gene content and gene order with relocation of most tRNA genes (Attardi, 1985; Milbury and Gaffney, 2005). At the species level, the extent of genome rearrangement between the two distinct gender-associated mitochondrial genomes appears to vary greatly among the three divergent bivalve lineages. In the Mytiloida, the gene order and content of F and M genomes from a species are conserved, but both lack the gene for ATPase subunit 8 (*atp8*) and have a second tRNA gene for methionine (*trnM*) (Mizi *et al.*, 2005; Breton *et al.*, 2006; Zbawicka, Burzynski and Wenne, 2007). By contrast, in the marine clam *Venerupis philippinarum*, gene content differs between M and F genomes as we observe a gene duplication for the cytochrome c oxidase subunit II gene (*cox2*) in the F genome and an extra *trnM* in the M genome, and both genomes have a short *atp8* gene (37 amino acids) the function of which is unclear (Attardi, 1985; Dreyer and Steiner, 2006). In freshwater mussels, an M genome-specific 3' extension of the

cytochrome c oxidase subunit II gene (Mcox2) has been clearly demonstrated (Curole and Kocher, 2002; Curole and Kocher, 2005; Walker *et al.*, 2006). This functional extension is a unique feature of unionoid M genomes and typically yields an ~80% increase in gene length relative to the female-transmitted cox2 gene (Chapman *et al.*, 2008). In unionoid bivalves, the presence of atp8 has been confirmed in both M and F genomes (Attardi, 1985; Serb and Lydeard, 2003). Also, the F genomes of *Inversidens japonensis* and *Hyriopsis cumingii* (both in the subfamily Gonideinae) exhibit a different gene order compared with the M genome of *I. japonensis* and the F genome of *Lampsilis ornata* [Serb and Lydeard, (2003); Zheng and Li, personal communication] (see Table 3.1). Analysis of gene order for additional F and M mtDNA genomes from freshwater mussels will allow us to test the following alternative hypotheses: (1) the translocation of several genes as observed in the F genomes of *I. japonensis* and *H. cumingii* represents an idiosyncratic gene rearrangement unique to these species or to the subfamily Gonideinae or (2) that this gene arrangement is in fact a characteristic shared with other unionoid species' F genomes or other subfamilies.

Table 3.1. Species and GenBank accession numbers of the sequences used in this study for phylogenetic analyses.

Species	Gender	GenBank accession Numbers
Bivalvia		
Autolamellibranchiata		
Unionoida		
<i>Venustaconcha ellipsiformis</i>	F	FJ809753
<i>Venustaconcha ellipsiformis</i>	M	FJ809752
<i>Pyganodon grandis</i>	F	FJ809754
<i>Pyganodon grandis</i>	M	FJ809755
<i>Inversidens japonensis</i>	F	AB055625
<i>Inversidens japonensis</i>	M	AB055624
<i>Quadrula quadrula</i>	F	FJ809750
<i>Quadrula quadrula</i>	M	FJ809751
<i>Lampsilis ornata</i>	F	NC_005335
<i>Hyriopsis cumingii</i>	F	NC_011763
Veneroida		
<i>Venerupis philippinarum</i>	F	NC_003354
<i>Venerupis philippinarum</i>	M	AB065374
<i>Acanthocardia tuberculata</i>	—	NC_008452
<i>Sinonovacula constricta</i>	—	NC_011075
<i>Hiatella arctica</i>	—	NC_008451
Pteriomorpha		
Mytiloida		
<i>Mytilus trossulus</i>	F	DQ198231
<i>Mytilus trossulus</i>	M	DQ198225
<i>Mytilus edulis</i>	F	NC_006161
<i>Mytilus edulis</i>	M	AY823624
<i>Mytilus galloprovincialis</i>	F	NC_006886
<i>Mytilus galloprovincialis</i>	M	AY363687
Ostreoida		
<i>Crassostrea hongkongensis</i>	—	NC_011518
<i>Crassostrea gigas</i>	—	NC_001276
<i>Crassostrea virginica</i>	—	NC_007175
Pectinoida	—	NC_009687
<i>Argopecten irradians</i>		
<i>Placopecten magellanicus</i>	—	NC_007234
<i>Mizuhopecten yessoensis</i>	—	NC_009081
<i>Chlamys farreri</i>	—	NC_012138
<i>Mimachlamys nobilis</i>	—	NC_011608
Out-group		
<i>Aplysia californica</i> (Gastropoda)	—	NC_005827
<i>Graptacme eborea</i> (Scaphopoda)	—	NC_006162
<i>Katharina tunicata</i> (Polyplacophora)	—	NC_001636
<i>Octopus vulgaris</i> (Cephalopoda)	—	NC_006353
<i>Platynereis dumerii</i> (Polychaeta)	—	NC_000931

Having additional unionoid F and M genomes available for comparative analyses would also significantly illuminate investigations into the likely unique origin of DUI (Theologidis *et al.*, 2008). Analyses of morphological and molecular datasets indicate that unionoid bivalves, together with trigonioid bivalves, compose a monophyletic taxon, the Paleoheterodonta (Healy, 1989; Hoeh *et al.*, 1998; Graf and Foighil, 2000). The relative antiquity of this taxon within the Bivalvia is supported by the molecular sequence-based phylogenies presented in Hoeh *et al.* (1998) [Hoeh *et al.*, (1998): Figure three; Giribet and Wheeler (2002): Figure five; Giribet and Distel (2003): Figure three point five and Dreyer and Steiner (2006): Figure five], which suggest that the Paleoheterodonta diverged early during the autolamellibranch (~suspension-feeding) bivalves phylogeny. Given the hypothesized relatively basal position of unionoids in the bivalve phylogeny, their mitochondrial genomes could retain plesiomorphic character states that are informative with respect to the initial mt genome duplication event (i.e., the formation of a distinct male-transmitted lineage in addition to a female-transmitted lineage) that led to the evolution of DUI.

In the present study, six new complete mitochondrial genomes, namely, the F and M genomes of the unionoid bivalves *Venustaconcha ellipsiformis* (Unionoida: Unionidae: Ambleminae: Lampsilini), *Pyganodon grandis* (Unionoida: Unionidae: Unioninae: Anodontini) and *Quadrula quadrula* (Unionoida: Unionidae: Ambleminae: Quadrulini), were compared with the available complete genomes of DUI species deposited in GenBank and their gene order, gene content and variation were analyzed. Additionally, complete bivalve mt genomes were analyzed to further test the hypothesized basal position of the Paleoheterodonta among extant autolamellibranch bivalves and to evaluate the evolutionary history of mitogenomic character state

transformations. The aim is to provide a context for comparisons of mt genomes among DUI and non-DUI bivalve lineages, and ultimately to identify the gene region(s) involved in the manifestation of DUI. Such descriptions and comparisons will contribute to a more complete picture of the evolution not only of the DUI system per se, but also of the factors involved in the near universal presence of SMI in animals.

3.3 Results

3.3.1 Phylogenetic analysis

The 50% majority-rule codon-based BI tree (Figure 3.1), derived from using concatenated sequences of mitochondrial protein-coding genes, is well resolved and similar in topology to the best BI tree produced from analysis of amino acids as well as to the best nucleotide- and amino acid-based ML and parsimony trees (not shown). In summary, these trees clearly indicate that (1) the Bivalvia is monophyletic, (2) pteriomorph and veneroid bivalves are reciprocally monophyletic with unionoid bivalves being the sister group to Pteriomorphia+Veneroida, (3) the F and M clades in unionoids are reciprocally monophyletic and (4) branch lengths indicate the higher substitution rate of the unionoid M genome relative to that of the unionoid F genome. The ML reconstruction of the presence/absence of DUI (Figure 3.2A), using the MK1 model, unambiguously indicates three origins of DUI for the taxa included in this study.

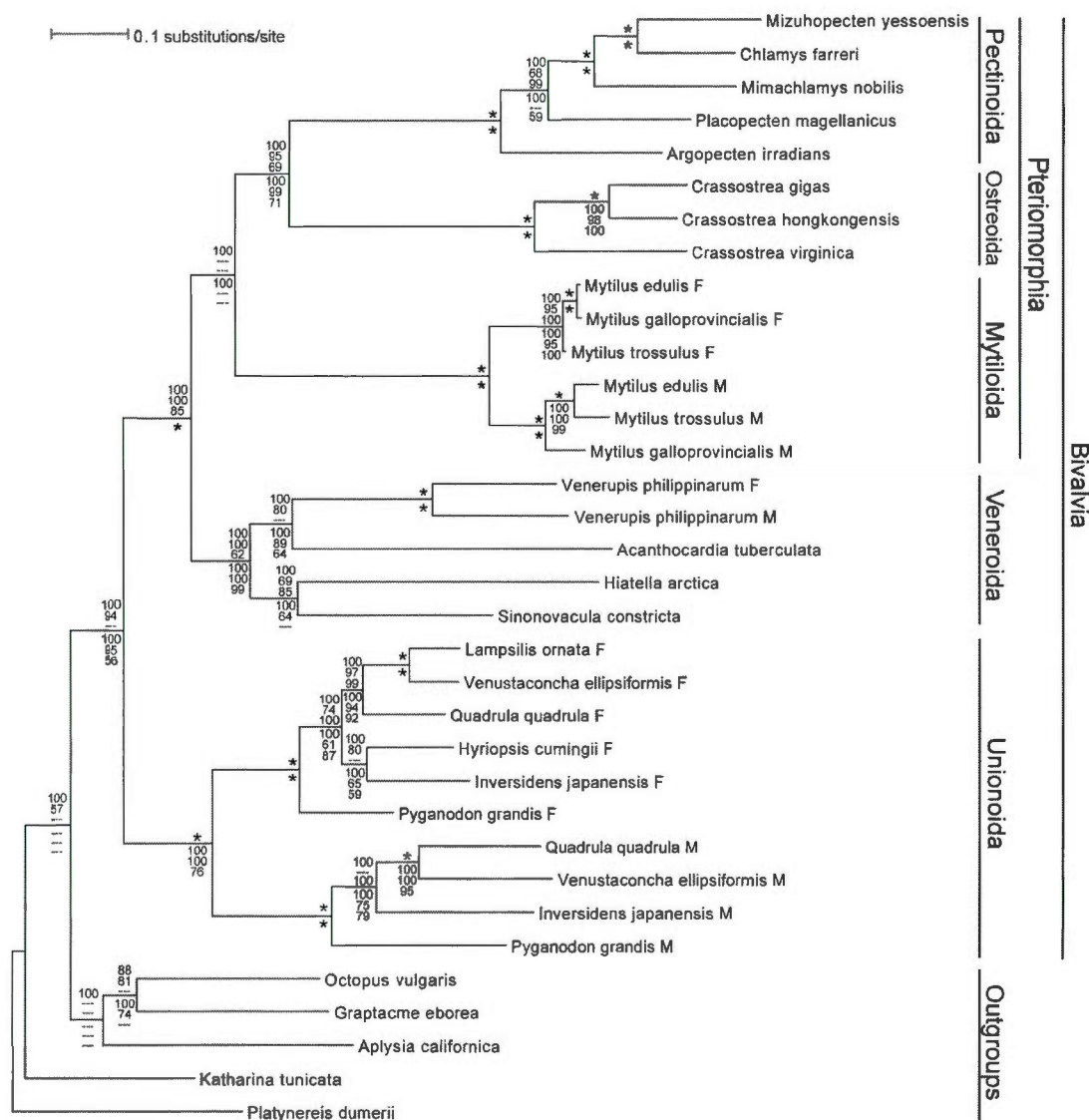


Figure 3.1 Bayesian inference majority-rule tree of bivalve mt genome relationships based on an analysis using the M3 codon substitution model and a nucleotide alignment of 12 mitochondrial protein-coding genes (atp8 excluded). Numbers above an internal branch, from top to bottom, indicate nodal support values from BI, ML and MP nucleotide-based analyses, respectively. Numbers below an internal branch, from top to bottom, indicate nodal support values from BI, ML and MP amino acid-based analyses, respectively. Only nodal support values > 50% are presented. An asterisk above an internal branch indicates that all three nucleotide-based nodal support values are 100; an asterisk below an internal branch indicates that all

three amino acid-based nodal support values are 100. Branch lengths reflect substitutions per site and the taxonomic and gender-specific transmission affiliations of the individual sequences are indicated at the right. All phylogenetic analyses strongly indicate that the unionoids represent the basal lineage for the bivalve taxa represented in this analysis.

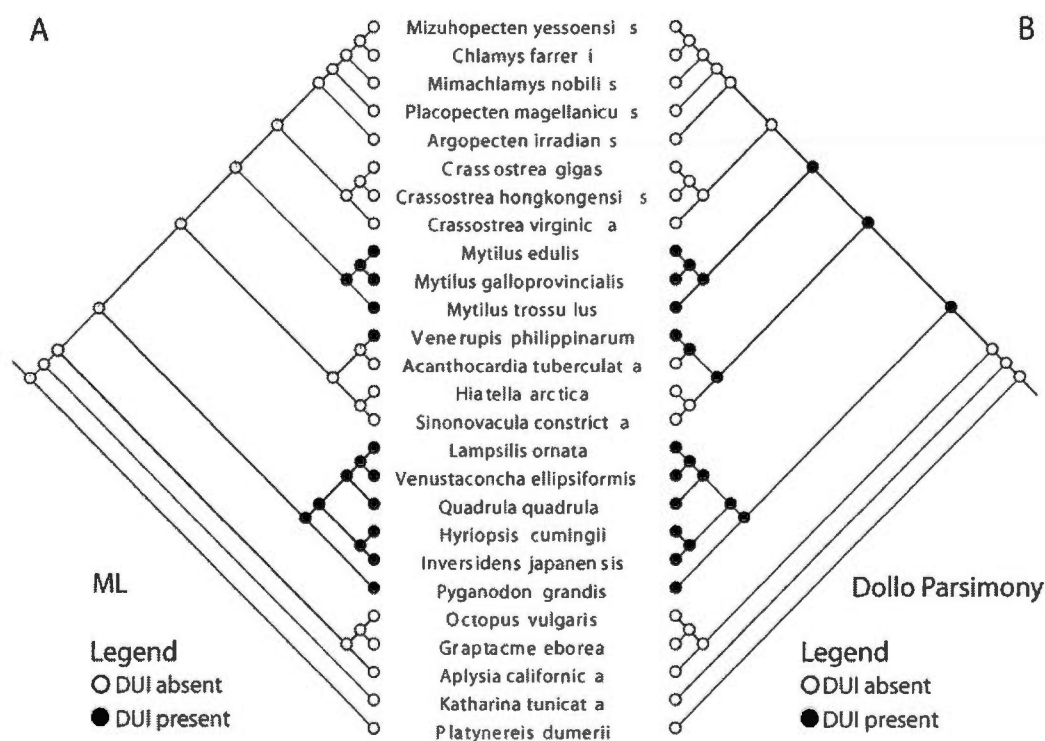


Figure 3.2 ML-based (A) and Dollo parsimony-based (B) ancestral character state reconstructions of presence/absence of DUI on a species-level tree derived from the tree presented in Figure 3.1. The ML-based reconstruction indicates three separate origins of DUI while the Dollo parsimony-based reconstruction indicates a single origin of DUI followed by three reversals to SMI.

3.3.2 Genome structural features

The length of the F *V. ellipsiformis*, F *P. grandis* and F *Q. quadrula* mitochondrial genomes are 15,975 bp, 15,848 bp and 16,033 bp, respectively. These sizes are ~950 bp to 1,200 bp smaller than their M counterparts whose genomes are 17,174 bp (*V. ellipsiformis*), 17,071 bp (*P. grandis*) and 16,970 bp (*Q. quadrula*). The length differences are mainly due to the presence of a unique M genome-specific 3' extension of the cytochrome c oxidase subunit 2 (cox2) gene and longer unassigned/noncoding regions in M genomes. The A+T composition is similar among the six newly sequenced genomes but higher than F and M *I. japonensis* (Table 2.2). Ribosomal RNA genes and protein-coding genes (except atp8) of M and F genomes of *V. ellipsiformis*, *P. grandis* and *Q. quadrula* are arranged identically but tRNA order differs among the analyzed genomes (Figure 2.3). In unionoids, ten or 11 genes out of ~37 are located on one strand and all the other genes on the opposite (Figure 3).

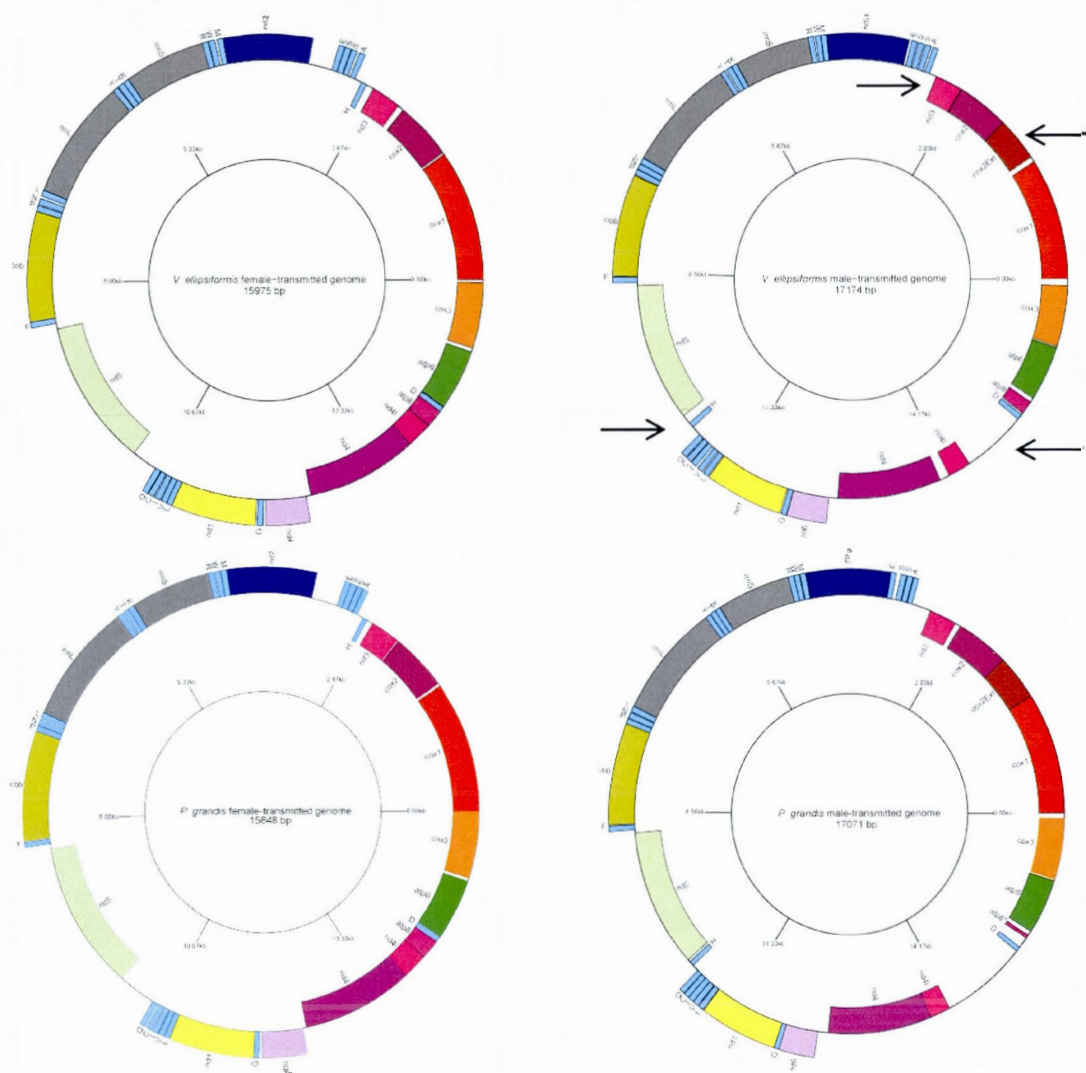


Figure 3.3 Gene maps of the M and F mitochondrial genomes of *Venustaconcha ellipsiformis*, *Pyganodon grandis*, *Inversidens japonensis* and *Quadrula quadrula*. Protein and rRNA genes are named as in the text while tRNA genes are abbreviated by the one-letter code of the corresponding amino acid [(L1 = trnL (cua), L2 = trnL (uaa), S1 = trnS (aga), and S2 = trnS (uaa)]. Genes positioned inside the plain line are encoded on the heavy strand and genes outside the line are encoded on the light strand. Atp8* (= genomes lacking full size atp8 gene). Black arrows on the *V. ellipsiformis* M genome indicate regions that differ between male- and female-transmitted genomes and the arrow on the *I. japonensis* F genome indicates the region

with a gene order distinct from that of the other figured F genomes. The circular gene maps of the genomes were drawn by GenomeVx (Conant and Wolfe, 2008) followed by manual modification.

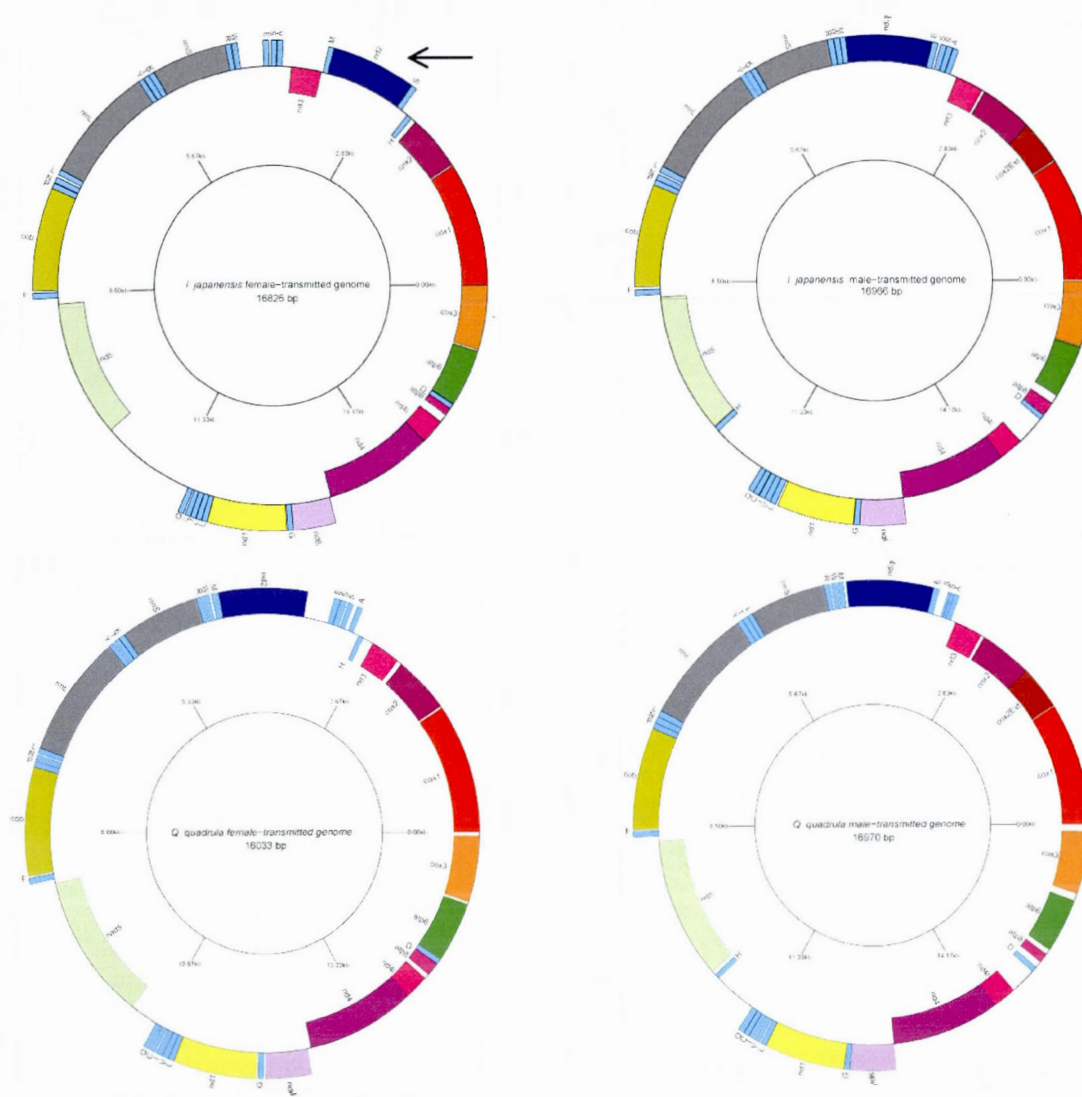


Figure 3.3 (continued).

Table 3.2. Main structural features of the female- and male-transmitted mitochondrial genomes of *Venustaconcha ellipsiformis*, *Pyganodon grandis*, *Inversidens japonensis* and *Quadrula quadrula*.

	<i>V. ellipsiformis</i> Female	<i>V. ellipsiformis</i> Male	<i>P. grandis</i> Female	<i>P. grandis</i> Male	<i>I. japonensis</i> Female	<i>I. japonensis</i> Male	<i>Q. quadrula</i> Female	<i>Q. quadrula</i> Male
Total size	15 975	17 174	15 848	17 071	16 826	16 966	16 033	16 970
pb								
A+T %	62.55	63.47	64.22	64.76	57.20	57.12	61.94	61.80
Strand – CG-skew	-0.46	-0.49	-0.36	-0.33	-0.42	-0.41	-0.48	-0.51
Strand – AT-skew	-0.40	-0.41	-0.30	-0.25	-0.41	-0.41	-0.44	-0.43
Strand + CG-skew	0.32	0.31	0.28	0.25	0.26	0.30	0.33	0.38
Strand + AT-skew	-0.06	0.04	-0.09	-0.08	0.00	0.05	-0.02	0.10
<i>rrnS</i>	865	859	847	841	844	864	857	865
<i>rrnL</i>	1271	1290	1271	1287	1304	1325	1297	1306
<i>cox1</i>	1542 (TTG/TAG)	1537 (GTG/T**)	1539 (TTG/TAA)	1537 (CTG/T**)	1539 (GTG/TAA)	1543 (TTG/T**)	1542 (TTG/TAG)	1534 (GTG/T**)
<i>cox2</i>	684 (ATG/TAA)	1233 (ATG/TAA)	681 (ATG/TAG)	1251 (ATG/TAG)	681 (ATG/TAA)	1224 (ATG/TAG)	681 (ATG/TAG)	1230 (ATG/TAG)
<i>Nd3</i>	357 (ATG/TAG)	354 (ATG/TAA)	357 (ATG/TAA)	357 (TTG/TAA)	381 (ATG/TAA)	360 (ATG/TAG)	360 (ATA/TAA)	366 (ATT/TAA)
<i>Nd2</i>	972 (ATG/TAG)	957 (ATT/TAA)	966 (ATG/TAG)	972 (ATT/TAA)	960 (ATG/TAA)	981 (ATA/TAG)	966 (ATG/TAA)	990 (ATC/TAA)
<i>cob</i>	1156 (ATT/T**)	1158 (ATG/TAA)	1146 (ATC/TAA)	1149 (TTG/TAA)	1152 (ATT/TAA)	1152 (GTG/TAA)	1152 (ATC/TAA)	1149 (ATG/TAA)
<i>Nd5</i>	1731 (ATA/TAA)	1758 (ATA/TAA)	1737 (GTG/TAG)	1740 (ATA/TAA)	1704 (GTG/TAG)	1765 (TTA/T**)	1734 (ATG/TAA)	1770 (ATG/TAA)
<i>Nd1</i>	903 (ATC/TAA)	903 (ATA/TAA)	900 (ATA/TAA)	897 (GTG/TAA)	897 (ATC/TAG)	879 (ATA/TAA)	894 (ATA/TAA)	912 (ATG/TAA)
<i>Nd6</i>	489 (ATT/TAA)	456 (ATA/TAG)	489 (ATC/TAA)	405 (ATA/TAA)	489 (ATC/TAA)	528 (ATA/TAG)	498 (ATG/TAA)	522 (ATA/TAG)
<i>Nd4</i>	1347 (ATT/TAG)	1350 (ATG/TAG)	1344 (ATT/TAA)	1347 (GTG/TAA)	1374 (GTG/TAA)	1362 (ATT/TAG)	1332 (TTT/TAA)	1380 (ATG/TAA)
<i>Nd4l</i>	297 (GTG/TAG)	297 (TTG/TAA)	297 (GTG/TAG)	264 (ATA/TAG)	297 (GTG/TAA)	309 (ATA/TAG)	294 (GTG/TAG)	297 (ATG/TAG)
<i>atp8</i>	210 (GTG/TAA)	174 (GTG/TAG)	225 (ATG/TAA)	Remnant	Remnant 96 (GTG/TAG)	181 (ATG/T**)	159 (ATG/TAA)	123 (GTG/TAA)
<i>atp6</i>	708 (ATG/TAG)	696 (ATG/TAA)	714 (ATG/TAG)	690 (ATG/TAA)	708 (ATG/TAG)	663 (ATG/TAG)	708 (ATG/TAG)	678 (ATA/TAG)
<i>cox3</i>	780 (ATG/TAG)	774 (ATT/TAA)	780 (ATG/TAA)	772 (ATG/T**)	795 (ATA/TAA)	810 (GTG/TAA)	780 (ATG/TAG)	777 (ATT/TAA)

3.3.2.1 tRNA Histidine

In the three newly sequenced M genomes of this study, tRNA histidine (trnH) is positioned between nd5 and nd1 whereas in F genomes, it is located between nd2 and nd3 (Figure 3.3). The location of the trnH gene in the *I. japonensis* M genome has been identified between cox1 and cox2 (Okazaki M and Ueshima R, personal communication). This location also corresponds to Mcox2e. However, our reannotation of the *I. japonensis* M genome identifies trnH between nd5 and nd1 as in the three new M genomes. Examination of 41 entire mt genomes across the Mollusca allow us to group some classes of mollusks according to the position of trnH. For example, nd5-trnH-nd4 (encoded on the heavy strand) is the common organization in the Cephalopoda while nd4-trnH-nd5 (encoded on the light strand) and cox2-trnG-trnH-trnQ-trnL2-atp8 (encoded on both strands) are two most common arrangements found in the Gastropoda. Except for unionoid genomes, no common arrangement is found in bivalves and the arrangement of the trnH-containing region appears to be unique to each genus or species sequenced to date. Nd2-trnR-trnH-nd4 is found in the genus *Crassostrea* spp. (oysters) whereas the marine mussels *Mytilus* spp. possess the arrangement nd2-trnR-trnW-trnA-trnS-trnH-trnP-nd3. Because of the current uncertainty regarding molluscan phylogeny, a rigorous ancestral character state reconstruction is not possible. However, the arrangement observed in *Mytilus* spp. could be an inversion (+ strand reversion) of the nd3-trnH-trnA-trnS1-trnS2-nd2 observed in the unionoid F genomes. In the four unionoid M genomes, trnH is located between nd5 and nd1 and the only other molluscan species with a similar location for its trnH (i.e., nd5-trnL-trnH-nd1) is the patellogastropod limpet *Lottia digitalis*. The gene order of the *L. digitalis* mt genome is the most

divergent among all gastropod mtDNAs sequenced thus far (Grande, Templado and Zardoya, 2008).

3.3.2.2. Extension of the M cytochrome c oxidase subunit II gene

The three analyzed M genomes possess the unique 3' extension of the cytochrome c oxidase subunit II gene (Mcox2e) (Curole and Kocher, 2002). In the three newly sequenced M genomes, the extension is 187 codons (*V. ellipsiformis* and *P. grandis*) or 186 codons (*Q. quadrula*) in length while the *I. japonensis* extension is slightly shorter with 181 codons.

3.3.2.3 Atp8 gene

As in the F genome of *L. ornata* (Serb and Lydeard, 2003) and the reannotated F and M genomes of *I. japonensis* (Gissi, Iannelli and Pesole, 2008), the newly sequenced F and M genomes contain the 13 protein-coding genes commonly found in other animal mtDNAs. Only the M genome of *P. grandis* appears to lack a complete atp8. In this species, a remnant of the atp8 gene that corresponds to the first 15 amino acids (MPQLSPVYWVSIFFL) of the protein, and that shows similarities with other atp8 genes sequenced in this study, has been identified between trnD and atp6 (Figure 3.3). Those 15 amino acids are followed by a stop codon. After the stop codon, we also identified an open reading frame, in a different frame than the first 15 amino acids, which could correspond to the remainder of atp8.

3.3.2.4 NADH dehydrogenase subunits 4 and 4L genes

Most unionoid mt genomes examined in this study have an overlap of 7 bp for subunits 4 and 4L of the NADH dehydrogenase complex (nd4 and

nd4l). Two exceptions are the M genome of *V. ellipsiformis*, which contains a noncoding region of 120 bp between nd4 and nd4l, and the F genome of *I. japonensis*, which possesses one nucleotide between those two genes.

3.3.3. Base composition and codon usage

The base composition bias of an individual strand can be described by skewness (Perna and Kocher, 1995), where $CG\text{-skew} = (C - G)/(C + G)$ and $AT\text{-skew} = (A - T)/(A + T)$. The strand encoding most of the proteins (including *cox1*) from the F and M genomes of all unionoid species has strong negative CG- and AT-skews (Table 3.2). Skews calculated for the opposite strand in all six genomes indicate complementary strand bias, with positive CG- and AT-skew values (Table 3.2), an expected result since, for example, A-skew on one strand is usually balanced by T-skew on the other (Francino and Ochman, 1997).

3.3.4 Transfer RNA genes

In all eight unionoid mt genomes, we identified all 22 tRNA genes according to their secondary structure features and their corresponding anticodons. Most have the potential to fold into a normal cloverleaf structure, although some do not have paired DHU arms, and a few others have mismatched bp. The putative cloverleaf secondary structures of unionoid tRNAs are available in the additional files (see appendices A.1 to A.8 for Supplementary Figures). The tRNA genes are ~60-70 bp long and the mean GC content varies between 35.2% and 37.4%. In the eight mt genomes, most of the tRNA genes are located on the light strand; only *trnH* (Histidine) and *trnD* (Aspartate) are located on the heavy strand along with most of the

protein-coding genes. As specified earlier, trnH has distinct localizations in F and M unionoid genomes (Figure 3.3). For all mt genomes (except M *I. japonensis*), the DHU arm of trnS1 (Serine) is unpaired. Unpaired DHU arms are also observed for the second Serine trnS2 (tct) and Threonine trnT in the M genome of *P. grandis*, for the Arginine trnR and Threonine trnT in the M genome of *V. ellipsiformis* and for the Cysteine trnC and Threonine trnT in the M genome of *Q. quadrula*. DHU arm for the Lysine trnK in the F genome of *Q. quadrula* is also unpaired. No unpaired DHU arm has been observed in tRNAs of M *I. japonensis*.

3.3.5 Unassigned regions and putative control regions

Twenty-two or thirty-three unassigned regions were detected in the six genomes, with sizes ranging from 1 to 1196 bp. The three newly sequenced F genomes present a more compact arrangement than the three M genomes (Table 2.3). We observed the opposite trend in *I. japonensis* where the F genome contains more unassigned sequences (12.9% of the genome) than the M genome (8.9% of the genome). The abundance of unassigned sequences in both F and M genomes of the four unionoid species analyzed here is similar with the results observed in their mytiloid F and M counterparts (~10% of unassigned sequences) (Mizi *et al.*, 2005; Breton *et al.*, 2006; Zbawicka, Burzynski and Wenne, 2007). In comparison, the veneroid clam *V. philippinarum* M and F genomes have a higher proportion of unassigned sequences (i.e., > 15.8% for F and > 21.3% for M) (Okazaki M, Ueshima R, personal communication).

Table 3.3. Unassigned regions of the female- and male-transmitted mitochondrial genomes of *Venustaconcha ellipsiformis*, *Pyganodon grandis*, *Inversidens japonensis* and *Quadrula quadrula*.

	<i>V. ellipsiformis</i> Female	<i>V. ellipsiformis</i> Male	<i>P. grandis</i> Female	<i>P. grandis</i> Male	<i>I. japonensis</i> Female	<i>I. japonensis</i> Male	<i>Q. quadrula</i> Female	<i>Q. quadrula</i> Male
Unassigned regions	27	33	24	27	27	22	29	28
Total (bp)	1310	1828	1178	2025	2177	1516	1366	1583
Proportion of the genome (%)	8.2	10.6	7.3	11.9	12.9	8.9	8.5	9.3
Largest NC region (bp)	308	833	480	1103	1196	698	346	555

3.3.6 Levels of intra- and interspecies sequence divergences

Table 3.4 contains intra- and interspecific comparisons of the concatenated nucleotide and amino acid sequences of 12 protein-coding genes (atp8 has been excluded and only species for which both F and M genomes are available were used) from the M- and F-transmitted mitochondrial genomes of unionoid bivalves. The smallest nucleotide and amino acid distances are observed between the two F genomes of the most closely-related species, i.e. *V. ellipsiformis* and *Q. quadrula* (P of 0.207 and P(aa) of 0.157). Overall, nucleotide and amino acid sequence divergences between all pairs of F genomes are considerably lower than between M pairs or between the M and F genomes of the same species. Early estimations of the nucleotide divergence between M and F genomes of unionoid bivalves were based on p-distances of partial cox1 sequences (Hoeh *et al.*, 1996a) and were around 28 to 33%. Our results show that these early estimations were conservative, and the average uncorrected divergence between the F and M concatenated nucleotide sequences for the 12 mitochondrial protein-

coding genes is 41% for *V. ellipsiformis*, 42% for *Q. quadrula* and 43% for *P. grandis* and *I. japonensis*. These very high intraspecific divergences are observed at both the DNA and protein levels (Table 3.4).

Table 3.4. Intra- and interspecies comparisons for 12 mitochondrial protein-coding genes: average pairwise sequence divergence

Pairs	P	K2P	JC	TrN	Na	P(aa)	D	K _a	K _s	K _a /K _s
F x F										
VenF x PygF	0.248	0.306	0.301	0.311	434	0.192	0.214	0.137	1.400	0.098
VenF x InvF	0.249	0.307	0.302	0.312	431	0.191	0.212	0.137	1.358	0.101
VenF x QuaF	0.207	0.246	0.242	0.251	353	0.157	0.170	0.101	1.010	0.100
PygF x InvF	0.274	0.346	0.340	0.353	488	0.216	0.244	0.157	1.785	0.088
PygF x QuaF	0.252	0.312	0.317	0.318	466	0.207	0.231	0.147	1.350	0.109
InvF x QuaF	0.247	0.303	0.308	0.308	466	0.207	0.231	0.144	1.213	0.119
M x M										
VenM x PygM	0.370	0.514	0.510	0.529	950	0.421	0.496	0.353	1.548	0.228
VenM x InvM	0.340	0.459	0.454	0.468	865	0.384	0.484	0.300	1.352	0.222
VenM x QuaM	0.303	0.393	0.389	0.402	796	0.353	0.435	0.272	0.946	0.288
PygM x InvM	0.372	0.522	0.514	0.533	882	0.391	0.496	0.327	2.093	0.156
PygM x QuaM	0.374	0.522	0.518	0.536	933	0.414	0.534	0.360	1.523	0.236
InvM x QuaM	0.340	0.458	0.454	0.465	836	0.371	0.463	0.307	1.241	0.247
F x M										
VenF x VenM	0.407	0.586	0.586	0.603	1135	0.503	0.700	0.474	1.112	0.426
PygF x PygM	0.428	0.636	0.634	0.657	1142	0.506	0.706	0.494	1.428	0.346
InvF x InvM	0.430	0.641	0.638	0.653	1139	0.505	0.703	0.482	1.523	0.316
QuaF x QuaM	0.419	0.625	0.614	0.632	1164	0.516	0.726	0.500	1.142	0.438

NOTE.—P, uncorrected nucleotide divergence; K2P, nucleotide divergence with Kimura's two-parameter model; JC, nucleotide divergence with Jukes and Cantor model; TrN, nucleotide divergence with Tamura-Nei model; Na, total number of amino acid differences; P(aa), uncorrected amino acid p-distances; D, Poisson-corrected amino acid distances, and estimates of K_a (the number of nonsynonymous substitutions per nonsynonymous site) and K_s (the number of synonymous substitutions per synonymous site) using the Jukes and Cantor (1969) correction. VenF: *Venustaconcha ellipsiformis* F genome; PygF: *Pyganodon grandis* F genome; InvF: *Inversidens japonensis* F genome; QuaF: *Quadrula quadrula* F genome; VenM: *V. ellipsiformis* M genome; PygM: *P. grandis* M genome; InvM: *I. japonensis* M genome; QuaM: *Q. quadrula* M genome.

3.4 Discussion

3.4.1 Phylogenetic analysis

The seemingly anomalous difference in branching pattern between the M and F *Mytilus* genomes is due to an asymmetric introgression of *M. edulis* M mtDNA into the Baltic *M. trossulus* (Zbawicka, Burzynski and Wenne, 2007). Nevertheless, the ingroup (= bivalves) topology in Figure 3.1 is consistent with other sequence-based phylogenetic hypotheses trees in that the Unionoida is basal to Pteriomorphia+Veneroida [e.g., Hoeh *et al.* (1998); Giribet and Distel (2003); Dreyer and Steiner (2006)] thus reinforcing the hypothesis that the Unionoida is a relatively ancient bivalve lineage potentially harboring the ancestral characteristics of DUI.

Although the phylogenetic hypothesis [[Pteriomorphia + Veneroida] Unionoida] is not typically portrayed as the best estimate of evolutionary relationships for these lineages at this time [e.g., Giribet (2008): Figure six point height], the statistical robustness of our phylogenetic analyses with regards to the Bivalvia (Figure 3.1) indicates that it should be seriously evaluated in future, higher level bivalve phylogenetic studies.

The three origins of DUI for the taxa included in this study (Figure 3.2A) runs counter to the prevailing hypothesis of a single origin for this complex trait with subsequent reversals to SMI (Zouros, 2000; Hoeh, Stewart and Guttman, 2002; Obata *et al.*, 2008; Theologidis *et al.*, 2008) but it is not unexpected given the bias toward "DUI absence" stemming from the difficulties in confirming the presence of DUI [e.g., Hoeh *et al.*, 1996b; Krebs, 2004; Walker *et al.*, 2006; Theologidis *et al.*, 2008]. The complexity of the cyto-nuclear interactions involved in DUI, and its very narrow taxonomic

distribution are consistent with the hypothesis that the gain of DUI is a relatively rare event with subsequent losses being potentially more common. If a low ratio of rate of DUI gain to rate of DUI loss actually holds, then the use of the Dollo parsimony model (Le Quesne, 1974; Farris, 1977) is more appropriate than the use of the MK1 model and the former indicates a single gain of DUI with three subsequent losses (Figure 3.2B). A much more accurate understanding of the actual taxonomic distribution of DUI combined with a taxonomically expanded version of our robustly supported bivalve phylogeny (Figure 3.1) would allow a rigorous evaluation of the single vs. multiple origins of DUI hypotheses as well as the rates of DUI gain vs. loss.

3.4.2 Genome structural features

Overall, the most notable differences observed between M and F unionoid genomes are (i) the position of *trnH*, (ii) an inversion of the *trnD* and *atp8* genes, (iii) the length of the *cox2* gene (the M genomes possess a 3' extension of *cox2*) as well as (iv) a noncoding region between *nd4* and *nd4l* in the M mtDNA genome of *V. ellipsiformis* (Figure 3.3). The female-transmitted mtDNAs of *V. ellipsiformis*, *P. grandis* and *Q. quadrula* are comparable in many respects to the F mtDNA of *L. ornata*, which is unique in gene arrangement relative to all other molluscan and metazoan mt genomes (Serb and Lydeard, 2003). The mitochondrial gene order rearrangement in the F genomes of *I. japonensis* and *H. cumingii* (Unionoida: Unionidae: Ambleminae: Gonideini), i.e., the relative positions of the *nd2*, *trnM* to *nd3* genes, appear to be unique to the Gonideini as neither *L. ornata*, *V. ellipsiformis* (Unionidae: Unionidae: Ambleminae: Lampsilini), *Q. quadrula* (Unionoida: Unionidae: Ambleminae: Quadrulini) nor *P. grandis* (Unionidae: Unionidae: Unioninae: Anodontini) show this rearrangement. We suggest that this distinct gene order in the F genomes of *I. japonensis* and *H. cumingii*

resulted from a tandem duplication of the gene region followed by the deletion of segments of the duplicated gene region. Losses and gains of genes, gene rearrangements and unusually large amounts of duplicated or noncoding nucleotides are common in mollusk mitochondrial genomes (Boore, 2006; Knudsen *et al.*, 2006; Smith and Snyder, 2007).

When looking across other complete bivalve genomes, which include species from the orders Pectinoida, Ostreoida, Veneroida and Mytiloida [Organellar Genome Retrieval database OGRE; Jameson *et al.*, 2003], all genes are characteristically on the same strand (Figure 3.4A). Among the Bivalvia, only in unionoids are the genes transcribed in both directions. Robust phylogenies are necessary to compare bivalve mt genome arrangements in an evolutionary context and our understanding of the phylogenetic relationships among the major lineages within the Bivalvia and among molluscan classes is still limited and controversial (Giribet, 2010). Nonetheless, based on bivalve phylogenetic analyses presented in Hoeh *et al.* (1998), Giribet and Distel (2003), Dreyer and Steiner (2006) and herein (Figure 3.1), the "all-on-one-strand" phenotype likely represents a shared, derived characteristic that evolved once in the common ancestor of the Pteriomorphia and Veneroida with the unionoid model of genes on different strands representing the ancestral state for the Bivalvia (Figure 3.4A).

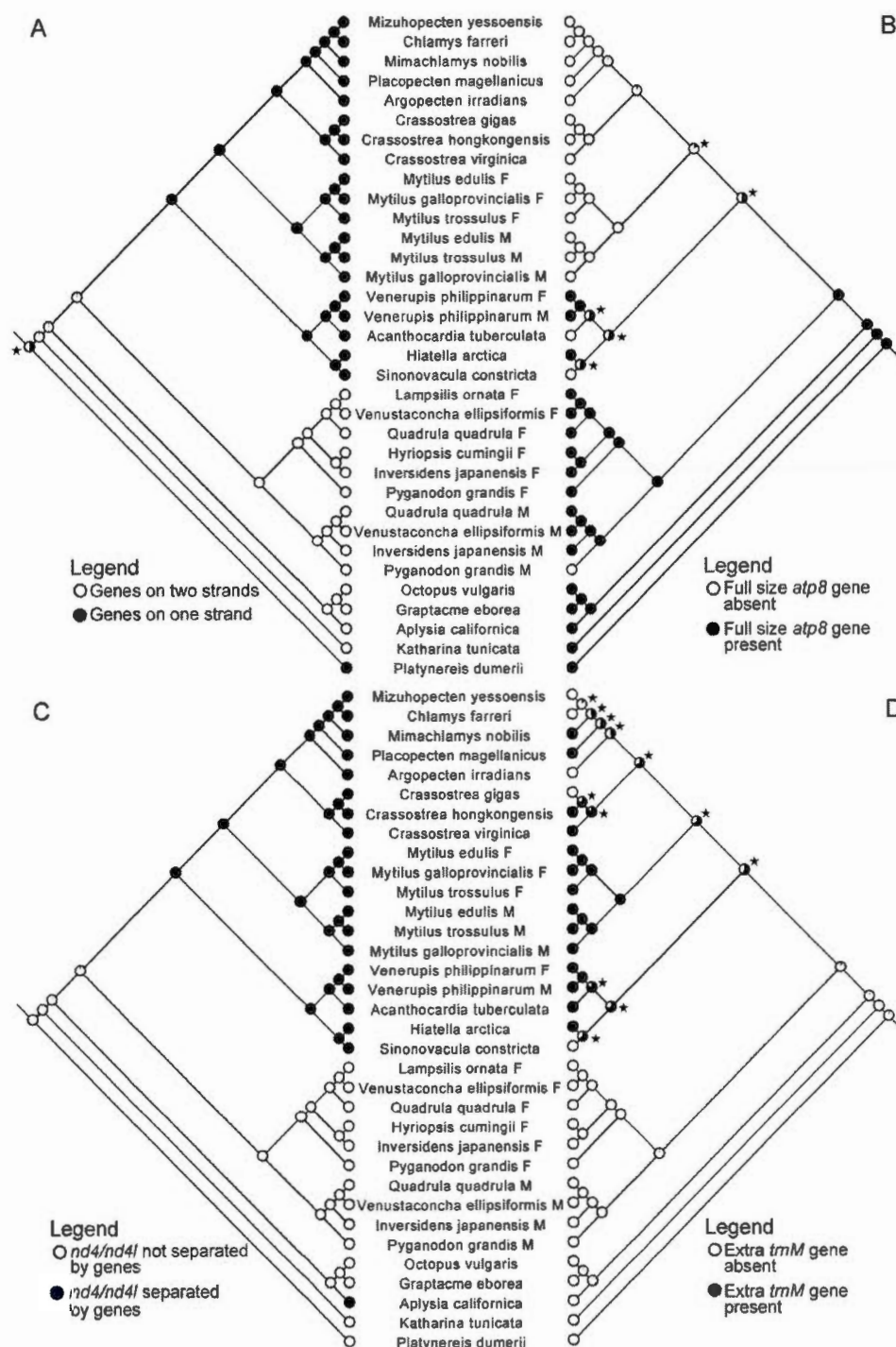


Figure 3.4. ML-based ancestral character state reconstructions of mt genomic characteristics on the tree presented in Figure 2.1: (A) genes on one strand vs. genes on two strands, (B) presence/absence of full size *atp8* gene, (C)

nd4/nd4l separated by genes vs. nd4/nd4l not separated by genes, (D) presence/absence of an extra trnM gene. For each of these four mt genomic characteristics, the indicated character state for the unionoid bivalve ancestor is identical to that indicated for the molluscan ancestor. "Stars" indicate nodes for which the character state designations were not statistically significant (all other nodes were statistically significant).

3.4.3 tRNA Histidine

While the gene boundary nd3-trH on the heavy strand observed in unionoid F genomes is not shared by any other mollusk taxon studied so far, the gene boundary nd5-trnH observed in the four unionoid M genomes is also shared by nine species of cephalopods, the polyplacophoran *Katharina tunicata* and the gastropod *Haliothis rubra*. This particular gene boundary could represent the ancestral character state for the Mollusca. The tRNA genes are the most evolutionarily mobile elements of the animal mitochondrial genome and variation in mitochondrial tRNA gene organization have been found in multiple divergent taxa (Boore, Lavrov and Brown, 1998; Downton and Austin, 1999). Rearrangement of tRNAs occurs frequently because their secondary structure facilitates their translocation (Mueller and Boore, 2005); alternatively, rearrangements can also result from a duplication event (Cantatore *et al.*, 1987).

3.4.4 Extension of the M cytochrome c oxidase subunit II gene

All unionoid bivalve M genomes examined to date contain an Mcox2e region (Walker *et al.*, 2006), which is not present in other DUI-possessing bivalve lineages nor, apparently, in any other animal mitochondrial genomes (Chapman *et al.*, 2008). Structural characterization of the MCOX2e region predicted the presence of an interspecifically variable number of

transmembrane helices (Chapman *et al.*, 2008), and immunohistochemistry- and immunoelectronmicroscopy-based analyses revealed that MCOX2e is expressed in sperm mitochondria (Chakrabarti *et al.*, 2006) and is sub-cellularly localized to both inner and outer mitochondrial membranes (Chakrabarti *et al.*, 2007). The latter localization, which possibly "tags" the outer surface of unionoid M genome-bearing mitochondria, could facilitate the differential segregation of the M genome-containing mt, derived from the fertilizing sperm, in male and female embryos [as observed in *Mytilus*; Cao, Kenchington and Zouros, 2004; Cogswell, Kenchington and Zouros, 2006]. Consistent with the above, seasonal variation in expression profiles suggest that unionoid MCOX2e functions in reproduction (Cantatore *et al.*, 1987; Chakrabarti *et al.*, 2007; Chapman *et al.*, 2008).

3.4.5 Atp8 gene

In animal mtDNAs, the *atp8* gene is the smallest protein-coding gene (\approx 50 to 65 aa) with only a few highly conserved amino acid residues. It encodes a protein subunit of the F₀ portion of the mitochondrial ATP synthase, which is the enzymatic complex that drives the phosphorylation of ADP to ATP. The ATP synthase comprises the F₁ catalytic domain situated in the mitochondrial matrix and the F₀ proton pore embedded in the mitochondrial inner membrane. Although the specific function of ATP8 is not yet known, in yeast, it is thought to play an important role in the assembly of the F₀ portion of ATP synthase and in determining ATP synthase activity [reviewed in Devenish *et al.*, (2000)]. In mammals, it is the most rapidly evolving mitochondrial protein-coding gene (Pesole *et al.*, 1999). *Atp8* has been lost independently from the mt genomes of several lineages, including some bivalves [e.g., marine mussels possessing DUI (Mizi *et al.*, 2005; Breton *et al.*, 2006; Zbawicka, Burzynski and Wenne, 2007) and oysters (Yu *et al.*, 2008), secernentean

nematodes (Okimoto *et al.*, 1992), and platyhelminths (Le *et al.*, 2000)]. Interestingly, all other mollusk species (belonging to gastropods, cephalopods, polyplacophorans and scaphopods) studied to date possess an *atp8* gene (Attardi, 1985; Yu *et al.*, 2008). The F and M genomes of the marine clam *Venerupis philippinarum* also possess a short putative *atp8* gene [37 aa; Dreyer and Steiner, 2006] and a potential remnant of the *atp8* gene has been found in the eastern oyster *Crassostrea virginica* (Milbury and Gaffney, 2005). These observations and the phylogeny displayed in Figure 3.1 reinforce the hypothesis that unionoid mt genomes possess the molluscan ancestral character state (= the presence of *atp8*) and that two losses of *atp8* in veneroids and another in the common ancestor of the Pteriomorphia could have occurred during bivalve phylogeny (Figure 3.4B). In the M genome of *P. grandis*, even though we identified an open reading frame that corresponds to a portion of *atp8*, the complete stop codon early in the sequence could yield a non-functional protein. Further analysis will be necessary to confirm or refute the existence of a functional *atp8* in the M genome of this species. For now, the presence/absence of *atp8* seems extremely labile across bivalve taxa, but this phenomenon does not appear to be related to the presence/absence of doubly uniparental inheritance.

3.4.6 NADH dehydrogenase subunits 4 and 4L genes

NADH dehydrogenase subunits 4 and 4L genes generally overlap or are adjacent to one another in animal mt genomes (Wolstenholme, 1992; Boore, 2006). This is also the case for most unionoid mt genomes examined in this study except for the M genome of *V. ellipsiformis* where those genes are separated by a large noncoding region as well as for the F genome of *I. japonensis*, where both genes are separated by one nucleotide. We cannot

exclude the possibility that this single nucleotide in the F *I. japonensis* sequence represents a sequencing error.

In vertebrates nd4 and nd4l are transcribed as one bicistronic mRNA, and are therefore localized together (Wolstenholme, 1992). Moreover, in several mollusks [i.e., one scaphopod, some gastropods and all 12 cephalopods studied to date — Organellar Genome Retrieval database OGRE; Jameson *et al.*, 2003], these two genes are also adjacent to one another or overlap. However, in all other non-unionoid bivalve species studied to date (7 genera), nd4 and nd4l have several intervening coding genes [e.g., *Crassostera gigas* (Milbury and Gaffney, 2005) and *Hiatella arctica* (Dreyer and Steiner, 2006)]. Again, the overlap observed between nd4-nd4l in most of the new unionoid genomes analyzed herein, and the phylogeny displayed in Figure 3.1 support the hypothesis that unionoid mt genomes possess the molluscan ancestral character state and that the derived state, "intervening genes" occurred once in the common ancestor of pteriomorph and veneroid bivalves (Figure 3.4C).

3.4.7 Base composition and codon usage

Although the exact mechanisms responsible for creating CG- and AT-skews like those observed in this study are still poorly understood, it is most likely created by the biases in mutational pressure owing to differences in the time spent as single-stranded DNA during both transcription and replication (Reyes *et al.*, 1998). The negative CG- and AT-skews observed in the strand that encodes most of the proteins [i.e., cox1 - cox3 - atp6 - atp8 - nd4L - nd4 - nd5 - nd3 - cox2 (Mcox2e)] and that make it G+T rich is reflected in the use of synonymous codons (see appendice B.1 for Supplementary Table). This is particularly evident at the third codon positions of protein-coding genes where C- and A-ending codons are used less frequently. Overall, TTT (Phe), TTG

(Leu) and TTA (Leu) are the most frequent codons, a result consistent with other invertebrate mtDNAs (Mizi *et al.*, 2005). Except for the stop codons TAA and TAG, TGC, CGC and ACG are among the least used codons. Of these, CGC is also the least common codon in the mtDNA of other mollusks (Mizi *et al.*, 2005).

3.4.8 Transfer RNA genes

Interestingly, unionoid bivalves do not possess an extra trnM, a situation that is present in both pteriomorph and veneroid bivalves (Attardi, 1985; Mizi *et al.*, 2005; Breton *et al.*, 2006). The presence of an extra trnM within the latter two lineages could represent a character state that was (1) independently derived multiple times, (2) derived once with multiple independent secondary losses or (3) both derived and lost multiple times independently (Figure 2.4D). The absence of a second trnM in all other molluscan species studied to date (Gissi, Iannelli and Pesole, 2008) reinforces the hypothesis that unionoids likely possess the molluscan ancestral character state for this character (Figure 2.4D).

3.4.9 Unassigned regions and putative control regions

The presence of multiple unassigned regions is not uncommon in mollusk mitochondrial genomes (Fuller and Zouros, 1993; Rigaa, Monnerot and Sellos, 1995) and is usually suggestive of molecular rearrangements (Carapelli *et al.*, 2008). The large unassigned region located between nd5 and trnQ in F genomes, and between trnH and trnQ in M genomes has been identified as a potential heavy strand origin of replication (OH) (Breton *et al.*, 2009). Otherwise, F and M unionoid mitochondrial genomes appear to

contain multiple and potentially bidirectional OL control regions (Breton *et al.*, 2009).

3.4.10 Levels of intra- and interspecies sequence divergences

It should be stressed that the measured divergences between unionoid F and M genomes considerably surpass intra- or interspecies values reported in classical model systems used for the study of intergenomic co-evolution (Blier, Dufresne and Burton, 2001; Blier *et al.*, 2006; Breton *et al.*, 2007). From a nucleo-mitochondrial evolutionary perspective, the question of how male freshwater mussels can tolerate heteroplasmy characterized by such variability remains to be solved.

Among species with DUI, freshwater mussels exhibit the greatest nucleotide and amino acid divergences between their gender-associated mtDNAs. For example, the average uncorrected nucleotide divergence observed between the F and M concatenated sequences of the 12 mitochondrial protein-coding genes of the marine mussel *Mytilus edulis* is about 23% (Breton *et al.*, 2006). The smaller level of divergence observed between the M and F mtDNAs in *Mytilus* is likely associated with periodic "role-reversal" or "masculinization" events, which are characterized by an invasion of the male route of inheritance by an F-like genome that becomes transmitted through sperm as a standard M genome (Venetis *et al.*, 2006; Hoeh *et al.*, 1997; Saavedra, Reyero and Zouros, 1997; Zouros, 2000). Specifically, the F-like, "recently-masculinized" M genome is only significantly different from a standard *Mytilus* F genome in that it possesses a so-called "standard M genome control region" and as its name implies, it is paternally transmitted [see Breton *et al.* (2006) and Stewart *et al.* (2009) for details]. Therefore, such masculinization events reset to zero the level of

mitochondrial gene sequence divergence between the M and F genomes. Complete absence of masculinization events, for over 200 million years, can explain the considerably greater divergences between unionoid M and F mtDNAs (Curole and Kocher, 2002; Curole and Kocher, 2005; Walker *et al.*, 2006). It has been proposed that the unionoid M-specific extension of the cytochrome c oxidase subunit II gene represents such specialization of the unionoid M genome that recombination (i.e., the addition of an M type genome's control region to an F genome) leading to role reversals are no longer possible in this taxon (Stewart *et al.*, 2009).

According to our results, ~50% amino acid divergence between unionoid F and M genomes can be tolerated by a species' nuclear environment without any major disruption of cytonuclear co-adaptation or impairment of mitochondrial function. This level of divergence could hardly be explained only by relaxation of selective pressure induced by a loss of metabolic function of M mtDNA since two recent studies had clearly shown the importance of M mtDNA gene products on sperm performance in *Mytilus edulis* (Jha *et al.*, 2008; Breton, Stewart and Blier, 2009). Further characterization of the conserved versus radical amino acid changes in evolutionarily conserved or non-conserved positions of mitochondrial proteins will help to delineate the levels/types of divergence in mtDNA encoded peptides that can be tolerated by a species' nuclear genome.

3.5 Conclusions

The basal position of the Unionoida within the autolamellibranch bivalves (Figure 3.1) and the hypothesized single origin of DUI [Figure 3.2B; (Zouros, 2000; Hoeh, Stewart and Guttman, 2002; Obata *et al.* 2008; Theologidis *et al.*, 2008) suggest that (1) DUI arose in the ancestral autolamellibranch bivalve lineage and was subsequently lost in multiple descendant lineages and (2) the DUI characteristics observed in unionoid bivalves could more closely resemble the DUI ancestral condition. We described the general features of eight mitochondrial genomes from unionoid bivalve species with the doubly uniparental mode of mitochondrial inheritance and highlighted several unusual characteristics of the M genomes, compared to their female-transmitted counterparts, e.g., the presence of Mcox2e and a novel localization of trnH. Based on the concatenated nucleotide sequences of 12 mitochondrial protein-coding genes, we determined an uncorrected amino acid p-distance between the M and F genomes of ~50%. From a nucleomitochondrial functional perspective, the question of how male freshwater mussels can tolerate heteroplasmy characterized by such variability remains to be solved as does the function(s) of DUI. Finally, the presence of the Mcox2e is one important feature that distinguishes markedly, but not solely, the unionoid M from the F genomes, but also the unionoid M from all other DUI-possessing bivalves as well as all other metazoan mtDNAs. This suggests that it could have been a facilitator of the transition from SMI to DUI in the ancestral autolamellibranch (assuming a single origin of DUI) or ancestral unionoid (assuming multiple origins of DUI) lineage. If the former hypothesis is corroborated, Mcox2e was subsequently lost from the M genome in the ancestor of the Pteriomorphia+Veneroida. Irrespective of a single vs. multiple origins of DUI, the ancestral character state reconstructions

in Figure 3.4 imply that significant mitochondrial genomic reorganization occurred in the Bivalvia subsequent to the divergence of the unionoid lineage. Studying additional complete bivalve mt genomes will give us the best hope of unraveling the origin(s) and function(s) of DUI as well as the origins and consequences of the unique mt genomic variation in the Bivalvia.

3.6 Methods

3.6.1 DNA extraction, PCR amplification and sequencing

Mitochondrial DNA was extracted from one male and one female *Venustaconcha ellipsiformis* from Gladwin County, Michigan (USA), one male and one female *Pyganodon grandis* from the northern Radisson region of Québec (Canada) (48°18'24"N, 68°29'34"W) and one male and one female *Quadrula quadrula* from the White River, Indiana (USA). Microscopic examination of gonadal tissues was performed to determine mussel gender based on the presence of eggs or sperm/sperm morulae. Total genomic DNA extractions were performed on the tissues (gonadal tissue for the male and the female *P. grandis* and gonadal tissue for the males and mantle tissue for the females *V. ellipsiformis* and *Q. quadrula*) using a QIAGEN DNA Extraction Kit (QIAGEN Inc., Mississauga, Canada) following the manufacturer's protocol. Subsequently, each mitochondrial genome was amplified by long-PCR from the genomic DNA in two large overlapping regions using the F-specific primer FCOIIhFor (5'-GCC TTA TGG GGT TGA TAG GCG AGT TCT TGT GAG G-3') with the amblesmine-specific primer AmbI16SFor (5'-CTG GGT TTG CGA CCT CGA TGT TGG CTT AGG GAA A-3'), and the F-specific primer FCOI195R (5'-GCA TAA CAA TTT CAC ACA GGC CAA TCA TYA TWG GYA TNA CCA-3') with the amblesmine-specific AmbI16SRev (5'-TTT

CCC TAA GCC AAC ATC GAG GTC GCA AAC CCA G-3') for the F genomes of *V. ellipsiformis* and *Q. quadrula*, and the M-specific primer MCOIIh63F (5'-CAC GAC GTT GTA AAA CGA CTT TAT RCC TRT KTG TGT RGA RGC TGT-3') with AmbI16SFor, and the M-specific MCOI19R (5'-GGA TAA CAA TTT CAC GGG TCC CAA TAT CYT TAT GRT TAG T-3') with AmbI16SRev for the M genomes of *V. ellipsiformis* and *Q. quadrula*. Both the F and M genomes of these species resulted in two fragments of ~11 kb and ~5.5 kb, respectively. The F genome of *P. grandis* was also amplified in two large overlapping regions using the F-specific primer 5'FCOIPygR (5'-TGC CAR TAA CAA RTA YAA AGT A-3') with ND1R (5'-GCT ATT AGT AGG TCG TAT CG-3; Buhay *et al.*, 2002; Smith and Snyder, 2007) and the F-specific primer UNIOND3 (5'-AGH SCK TTT GAR TGY GKG TTT GA-3') with ND1F (5'-TGG CAG AAA AGT GCA TCA GAT TTA AGC-3; Palumbi, 1996; Smith and Snyder, 2007). We obtained two fragments of ~11 kb and ~7 kb, respectively. The M genome of *P. grandis* was amplified with PygMcox2eF (5'-TTG AAG CAG TTA GAG TTG AGG T-3') in combination with 16Sar-L2 (5'-CGC CTG TTT AYC AAA AAC AT-3'; modified from (Wyman, Jansen and Boore, 2004) and PygMcox2eR (5'-TAY AAT CTT YCA ATR TCY TTA TGA TT-3') combined with 16Sbr-H2 (5'-CCG GTC TGA ACT CAG ATC RYG T-3'; modified from Wyman, Jansen and Boore, 2004). PygMcox2eF and PygMcox2eR were specifically designed to amplify the M genome. The resulting fragments were about ~12 kb and ~5.5 kb.

Long-PCR amplifications were performed in 50 µl reaction volumes using the QIAGEN LongRange PCR Kit in similar conditions to the manufacturers' suggestions: 1× LongRange PCR Buffer with 2.5 mM of Mg²⁺, 500 µM of each dNTP, 0.4 µM each primer, 2 U of LongRange PCR Enzyme Mix and ~25 ng of template DNA. For the M and the F genomes of *V. ellipsiformis* and

Q. quadrula, reactions were cycled at 85°C for 60 s, 93°C for 60 s, and 37 cycles of 93°C for 15 s, 53°C for 30 s and 68°C for 6 min for the short fragment or 11 min for the longest one. Thermal cycling conditions for the M and F genomes of *P. grandis* were as follows: 93°C for 3 min, followed by 35 cycles of 93°C for 15 s, 46-54°C for 30 s and 68°C for 7-12 min and a final extension at 72°C for 10 min. Each amplicon appeared as one abundant band of the appropriate size on an agarose gel. The resulting PCR products were gel purified using QIAGEN QIAquick Gel Extraction Kit. Following DNA quantification for each amplicon, the two amplicons (~5 µg from each) for each genome were pooled and then processed for direct sequencing in a single reaction by the 454 Life Sciences Massively Parallel Pyrosequencing Platform (whole genome sequencing protocol) of the McGill University and Genome Quebec Innovation Center.

For the M and the F genomes of *V. ellipsiformis*, amplifications were pooled and a total of 10,413 reads were produced to provide at least 45× coverage of the complete mitochondrial genomes. The sequences were then assembled into a single contig of 15,975 base pairs (bp) for the F genome and 17,174 bp for the M genome. For the M and the F genomes of *Q. quadrula*, amplifications were pooled and a total of 11,978 reads were produced to provide at least 66× coverage of the complete mitochondrial genomes. The sequences were assembled into a single contig of 16,033 bp for the F genome and 16,970 bp for the M genome. For the *P. grandis* F genome, draft assemblies were based on 14,794 total reads. The initial assembly of the 454 pyrosequencing data into two predominant contigs (~6.7 kb) and a small one (834 bp) was provided by 454 Life Sciences (Branford, CT, USA), and corresponded to a mitochondrial genome coverage of 115× and 437× respectively. The final assembly in one large contig of 15,848 bp

was performed using SeqMan (DNASTar Inc., Madison, WI, USA). The complete M genome of *P. grandis* was generated from assembly of 7,652 successful sequence reads into a single contig of 17,071 bp which corresponded to an overall mitochondrial genome coverage of > 100x.

The complete sequences of the F and M mitochondrial genomes for *Venustaconcha ellipsiformis*, *Pyganodon grandis* and *Quadrula quadrula* can be accessed under the GenBank accession numbers FJ809753, FJ809752, FJ809754, FJ809755, FJ809750, and FJ809751, respectively.

3.6.2 Gene annotation and analysis

The complete F and M mitochondrial genomes for each species were initially analyzed with the NCBI Open Reading Frame Finder using the invertebrate mitochondrial code. Protein-coding and ribosomal RNA genes were annotated using DOGMA (Wyman, Jansen and Boore, 2004) and then aligned with the mtDNA genes annotated in GenBank using ClustalW (Thompson, Higgins and Gibson, 1994). The 5' and 3' ends of both *rrnL* and *rrnS* genes were assumed to be next to the ends of bordering tRNA genes. Mitochondrial tRNA genes were identified and confirmed using a combination of programs: tRNAscan-SE 1.21 (Lowe and Eddy, 1997) with a COVE cutoff score of 0.1, DOGMA (Wyman, Jansen and Boore, 2004) and ARWEN (Laslett and Canback, 2008). Mitochondrial gene order comparisons were facilitated using the OGRE web site at <http://drake.physics.mcmaster.ca/ogre/index.shtml> (Jameson *et al.*, 2003).

Basic sequence statistics and evolutionary distances among genes were performed using MEGA version 4.0 (Kumar *et al.*, 2001) and DnaSP version 4.0 (Rozas *et al.*, 2003). To estimate evolutionary distance between pairwise comparisons, the following parameters were used: uncorrected nucleotide divergence (P_i = uncorrected nucleotide diversity), nucleotide divergence using the Jukes-Cantor (JC), Kimura two-parameter (K2P), and Tamura and Nei (TrN) models of nucleotide substitution. Estimated parameters also included total amino acid differences (Na), uncorrected amino acid distances ($p(aa)$), poisson-corrected amino acid distances (D), number of synonymous substitutions per synonymous site (Ks) and number of nonsynonymous substitution per nonsynonymous site (Ka) (Nei and Gojobori, 1986). The Jukes-Cantor correction for multiple substitutions was applied. Strand asymmetry was measured using the formulas $AT\text{-skew} = (A - T)/(A + T)$ and $CG\text{-skew} = (C - G)/(C + G)$ (Perna and Kocher, 1995; Hassanin, Leger and Deutsch, 2005) and calculated with MEGA 4.0 (Kumar *et al.*, 2001) at fourfold redundant sites for each mitochondrial protein-coding gene.

Phylogenetic hypotheses for the Bivalvia, using Bayesian inference (BI), maximum likelihood (ML) and maximum parsimony (MP), were constructed using concatenated nucleotide and amino acid sequences from 12 protein-coding genes (we excluded *atp8* owing to alignment issues and its apparent absence in multiple bivalve species). We used both Clustal W (Larkin *et al.*, 2007) and Dialign version 2.2.1 (Subramanian, Kaufmann and Morgenstern, 2008) for the alignments, with subsequent manual adjustments, and the amino acid alignment was used as a template to align the corresponding codons. Amino acids from 29 complete bivalve mitochondrial genomes and those from five outgroup species (gastropod *Aplysia californica* [NC_005827], cephalopod *Octopus vulgaris* [NC_006353], scaphopod *Graptacme eborea*

[NC_006162], polyplacophoran, *Katharina tunicate* [NC_001636], polychaete *Platynereis dumerii* [NC_000931]) were aligned using Clustal W in MEGA 4.0 and manually reviewed (Table 3.1). Regions of ambiguous alignment were excluded prior to the phylogenetic analyses. The analyzed matrices had either 7,704 nucleotide positions or 2,568 amino acid positions, and these files are available from the authors.

The codon- and amino acid-based BI analyses were conducted with Mr. Bayes version 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The codon-based analysis invoked the M3 model (Yang *et al.*, 2000) with two simultaneous runs of 5 million generations each (a total of 50,000 saved trees/run). The amino acid-based BI analysis invoked the variable rate "glorified GTR model" [see the MrBayes manual; Ronquist, Huelsenbeck and van der Mark, 2005] with two simultaneous runs of 2.9 million generations each (a total of 29,000 saved trees/run). Both sets of BI analyses reached convergence (average standard deviation of the split frequencies was <0.01) and the burnin for each set was determined by reference to the log probability of observing the data \times generation plot (codon-based BI run burnin = 4 million generations [= the last 10,000 trees/run saved contributed to the majority-rule tree], amino acid-based BI run burnin = 1.9 million generations [= the last 10,000 trees/run saved contributed to the majority-rule tree]).

Codon and amino acid-based ML analyses were conducted with Garli version 0.96 (Zwickl, 2006). The M3 model was used in the codon-based ML analysis which was set to use the observed nucleotide frequencies at each codon position separately. A non-parametric bootstrap (Felsenstein, 1985) analysis was performed, using 300 replicates, to assess nodal support for the

codon analysis-based trees. The program ProtTest [<http://darwin.uvigo.es/software/prottest.html>] was used to evaluate the best amino acid model for our data from those models available in Garli. Both the Akaike information criterion and Bayesian information criterion selected the WAG+F model (Whelan and Goldman, 2001) as the best available model that was therefore used in the ML amino acid analysis. A non-parametric bootstrap was performed, using 600 replicates, to assess nodal support for the amino acid analysis-based trees.

Maximum parsimony analyses were conducted with PAUP* (Swofford, 2002). The nucleotide-based MP analysis utilized equally weighted transversion parsimony (= only purines vs. pyrimidines were coded) and 1000 random addition runs for estimating the most parsimonious tree. A non-parametric bootstrap transversion parsimony analysis was run (with 1000 replicates) using 10 random addition runs per replicate. The amino acid-based MP analysis was carried out with equal weighting and 1000 random addition runs were used to estimate the most parsimonious tree. Lastly, an equally weighted parsimony, non-parametric, bootstrap analysis was run on the amino acid matrix (with 1000 replicates) using 10 random addition runs per replicate.

The estimation of ancestral mitogenomic character states and the presence/absence of DUI, based on the majority-rule codon-based BI tree, was carried out using the ML algorithm in Mesquite version 2.6 (Maddison and Maddison, 2008). An estimation of ancestral character states for the presence/absence of DUI, using the best codon-based BI tree and Dollo parsimony, was done with MacClade version 4.07 (Maddison and Maddison, 2003). The asymmetry likelihood ratio test was used to determine whether the

AsymmMK model was significantly better than the MK1 model (see the Mesquite manual). The MK1 model was used in all likelihood reconstructions because in all cases, the AsymmMK model was not a significantly better model, therefore, we used the simpler model (the MK1 model has one less parameter). The use of a likelihood ratio test to calculate P-values for ancestral states is not possible because hypotheses regarding the likelihoods of each possible state at a given node are non-nested. Therefore, to make decisions regarding the significance of ancestral character states Pagel (1999) following Edwards (1972), recommended that ancestral character state estimates with a log likelihood two or more units lower than the best state estimate (decision threshold [T] set to $T = 2$) be rejected. Generally viewed as a conservative cutoff, this threshold has been used by numerous recent authors [e.g., Fernandez and Morris, 2007; Murphy *et al.*, 2007; Koepfli *et al.*, 2008]. For the data presented herein, this protocol ensures that all of the character states judged to be significant have proportional likelihoods at least 10 times greater than that of any other state.

3.7 Authors' contributions

HDB and SB conceived the study, sequenced and annotated the new mitochondrial genomes. HDB aligned the sequence, analyzed the data, and drafted the manuscript. EGC conducted the computational phylogenetic analysis. PUB participated in coordination of the study and also helped to improve the manuscript. AEB provided tissue samples for DNA extraction and material support for the sequencing. DTS helped in interpreting the data and revising the manuscript. WRH also conceived the study and participated in its design, in data analysis and in drafting the manuscript. All authors read and approved the final manuscript.

3.8 Acknowledgements

This work was supported by research grants from the National Science Foundation (NSF DEB-0237175 to W. R. Hoeh), the National Sciences and Engineering Research Council (NSERC to P.U. Blier and D.T. Stewart) and by A. E. Bogan's research funds. H. Doucet-Beaupré and S. Breton were financially supported by a NSERC scholarship and a NSERC fellowship, respectively. E. G. Chapman is supported by the Kentucky Agricultural Experiment Station State Project KY008043.

CHAPITRE IV

PYGANODON (BIVALVIA: UNIONOIDA: UNIONIDAE) PHYLOGENETICS: A MALE- AND FEMALE-TRANSMITTED mtDNA GENOMIC PERSPECTIVE

Hélène Doucet-Beaupré¹, Pierre U. Blier¹, Eric G. Chapman², Helen Piontkivska³, France Dufresne¹, Bernard E. Sietman⁴, Renee S. Mulcrone³ and Walter R. Hoeh³

¹Département de Biologie, Université du Québec à Rimouski, 300 Allée des Ursulines, Rimouski, Québec, Canada, G5L 3A1. (H. Doucet-Beaupré : conception du projet, collecte de spécimen, expérimentation et analyse des données, rédaction de l'article et conception des figures ; P. U. Blier : supervision et aide à la rédaction, F. Dufresne : soutien technique).

²Department of Entomology, University of Kentucky, S225 Agricultural Science Center N Lexington, Kentucky, 40546-0091, USA. (E.G. Chapman: aide au traitement des données et à la rédaction).

³Department of Biological Sciences, Kent State University, Kent, Ohio, 44242, USA. (H. Piontkivska: aide au traitement des données; R.S. Mulcrone soutien technique; W. R. Hoeh : aide au traitement des données et à la rédaction).

⁴ Minnesota Department of Natural Resources, Saint Paul, Minnesota, 55155, USA. (B. E. Sientman: collecte de spécimen)

Key Words: Unionoida, Phylogeography, Doubly uniparental inheritance, *Pyganodon*.

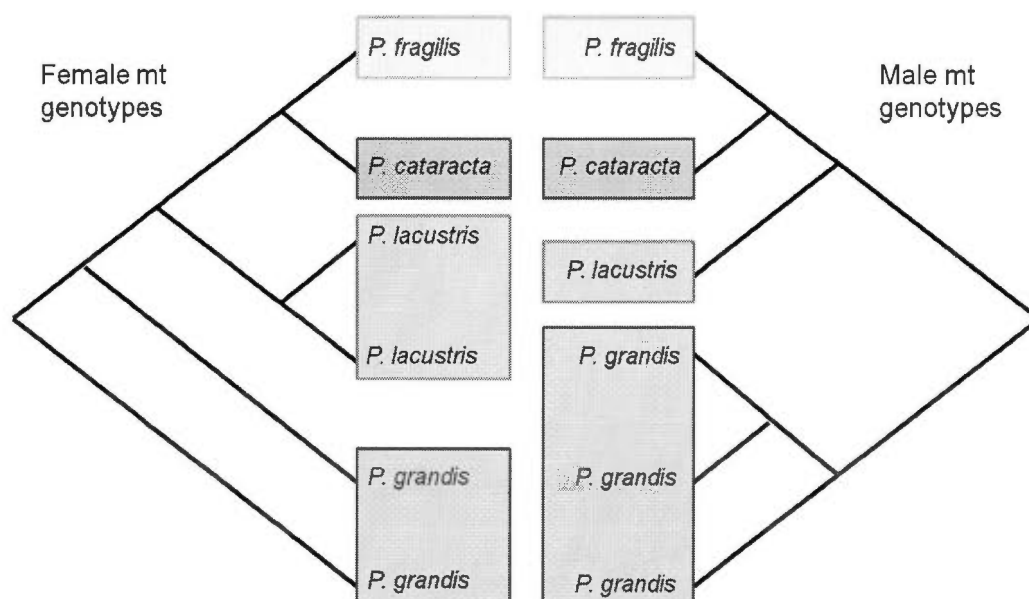
Running Head: Phylogenetics inferred by M and F mtDNA in *Pyganodon*

Résumé

La délimitation des espèces, les relations phylogénétiques et la répartition géographique de nombreuses espèces de bivalves unionoïdes, comme celles du genre *Pyganodon*, demeurent non résolues dans l'est de l'Amérique du Nord. La compréhension de la variation génétique au sein et entre les populations ainsi qu'entre les espèces sont cruciales pour la planification de stratégies efficaces de conservation de ces espèces. Qui plus est, les bivalves unionoïdes sont un des groupes animaux les plus menacés dans le monde. La conservation des espèces de ce groupe est indispensable du point de vue des habitats dulcicoles, mais également parce qu'elles possèdent un système unique d'hérédité mitochondriale où des lignées d'ADN mitochondrial associées sexes coexistent : un génome transmissible par les femelles (F) et un génome transmissible par les mâles (M). Dans cette étude, quarante-deux populations de *Pyganodon* spp. ont été échantillonnées sur une vaste étendue géographique et des fragments de deux gènes mitochondriaux (COX1 et COX2) ont été amplifiés à la fois dans les ADNmt M et F. Nos résultats confirment le caractère récent de la divergence entre *P. cataracta* et *P. fragilis*. Nous avons également trouvé deux lignées F et M relativement divergentes au sein de *P. grandis*. Les relations entre les lignées F et M des arbres de *P. grandis* ne sont pas congruentes. Nous avons trouvé qu'un haplotype de *P. lacustris* aurait récemment envahi tout l'espace génotypique M conduisant à une diversité étonnamment faible dans la lignée mitochondriale M chez cette espèce. Notre étude a soulevé des questions qui obligent à repenser les processus d'hybridation et de délimitation des espèces dans le genre *Pyganodon*. Comme les génomes mâle et femelle n'affichent pas toujours la même histoire phylogénétique dans chaque espèce, nous avons également discuté de l'importance d'être prudent dans l'interprétation des données moléculaires basées uniquement sur les génomes transmis de façon maternelle.

4.1 Abstract

Species boundaries, phylogenetic relationships and geographic distributions of many unionoid bivalve species, like those in the genus *Pyganodon*, remain unresolved in eastern North America. Because unionoid bivalves are one of the most imperiled groups of animals in the world, understanding the genetic variation within and among populations as well as among species is crucial for effective conservation planning. Conservation of unionoid species is indispensable from a freshwater habitat perspective but also because they possess a unique and model mitochondrial inheritance system where distinct gender-associated mitochondrial DNA lineages coexist: a female-transmitted (F) genome and a male-transmitted (M) genome. In this study, forty-two populations of *Pyganodon* spp. were sampled across a large geographical range and fragments of two mitochondrial genes (cox1 and cox2) were amplified from both the M- and F-transmitted mtDNA genomes. Our results support the recency of the divergence between *P. cataracta* and *P. fragilis*. We also found two relatively divergent F and M lineages within *P. grandis*. The relationships among the *P. grandis* specimens in the F and M sequence trees are not congruent. We found that a single haplotype in *P. lacustris* have recently swept throughout the M genotype space leading to an unexpectedly low diversity in the M lineage in that species. Our survey put forward some challenging results that force to rethink hybridization and species boundaries in the genus *Pyganodon*. As the male and the female genomes do not always display the same phylogenetic story in each species, we also discuss the importance of being careful in the interpretation of molecular data based solely on maternal transmitted mtDNA genomes.



4.2 Introduction

Unionoid bivalves, also called freshwater mussels, are an important but neglected component of freshwater aquatic ecosystem. They are long-lived partially infaunal filter feeders that can be found in most permanent freshwater bodies. North America boasts the richest unionoid fauna in the world with nearly 300 currently recognized species and subspecies, representing two families (Unionidae and Margaritiferidae) in the United States and Canada (Bogan, 1993, Turgeon *et al.*, 1998, 1988; Williams *et al.*, 1993). Unionid bivalves have a unique and complex life cycle. They are ovoviviparous and depend on host species for reproduction since their life cycle include a short obligate parasitic larval stage on fish or amphibian hosts (Bauer and Wächtler, 2000). As adults, they have limited dispersal abilities and the release of spermatozoa from males occurs in the open water of lakes, rivers and streams, which is the main dispersal mode along with the parasitic larval stage (Nagel, 2000). The insular nature and diversity of freshwater habitats (Strayer, 2006), their limited dispersal opportunities (Bilton, Freeland and Okamura, 2001), their high degree of variation and endemism and association with major drainage system with relatively long geological histories (Burch, 1975; Burky, 1983; McMahon, 1991) as well as their unique and complex combination of life history traits (Haag and Leann Staton, 2003) has deeply influenced the observed genetic diversity and geographic variation in unionid bivalves.

Unionid bivalves are also one of the most imperiled groups of animals in the world (Lydeard *et al.*, 2004), with 70% of the recognized species in North America considered extinct, endangered, threatened or of special concern (Williams *et al.*, 1993). Unionids are generally sensitive to ecosystem deterioration and have declined for various reasons such as watershed

alterations, changes in water quality resulting from organic enrichment and chemical contamination, and the recent introduction of the exotic zebra mussel, *Dreissena polymorpha* (Nalepa, 1994; Schloesser and Nalepa, 1994), the quagga mussel, *Dreissena bugensis*, and the Asian clam, *Corbicula fluminea* (McMahon, 1983; Ricciardi *et al.*, 1998). Effective conservation planning for unionoids includes protecting distinct species and populations and requires an understanding of genetic variation within and among populations for proper species and subspecies delineation. Unionoid species identification is usually based on conchological characteristics but phenotypic plasticity in shell morphology renders the differentiation of particular species problematic [e.g., species within the genus *Pyganodon*; Cyr *et al.*, 2007]. Alternatively, molecular markers can be used for a more accurate estimate of species diversity within unionoid lineages and to investigate population genetics and phylogeography.

Regardless of its unquestionable relevance to freshwater mussel conservation, the study of unionoid genetics is of pivotal importance in regard to this taxon's unique mitochondrial inheritance system. Unionoid bivalves possess a distinct form of mitochondrial DNA (mtDNA) inheritance, termed "doubly uniparental inheritance" [DUI; for review, see Breton *et al.* (2007); Passamonti and Ghiselli (2009)]. In these organisms, distinct gender-associated mitochondrial DNA lineages coexist: a female-transmitted (F) genome and a male-transmitted (M) genome. Under DUI, female bivalves transmit their mitochondria (carrying F mtDNA) to both sons and daughters, as in standard maternal inheritance, but males pass on their mitochondria (via sperm carrying M mtDNA) to only sons [e.g., Breton *et al.* 2007; but see Obata *et al.*, 2006; Chakrabarti *et al.*, 2007; Obata *et al.*, 2007]. At the organismal level, male bivalves with DUI are thus heteroplasmic and contain

both M and F genomes. Evidence has recently been presented that is consistent with the hypothesis that the DUI system in unionid and margaritiferid bivalves functions in the maintenance of separate sexes (= dioecy; Breton *et al.*, 2011). The bivalve order Unionoida is the only freshwater taxon where DUI is present; the two other bivalve lineages possessing DUI are marine (Mytiloida and Veneroida). This nontraditional and unique mode of mitochondrial inheritance provides the opportunity to study mitochondrial behavior from a different and privileged point of view and could be very useful in answering many systematic questions that cannot be effectively solved by studying taxa with the classical animal mitochondrial transmission mode (i.e., strict maternal inheritance [SMI]) (Passamonti and Ghiselli, 2009). Specifically, studies of intra- and interspecific genetic variation for both M and F genomes potentially offer complementary views on species limits, evolutionary relationships, and levels of population variation (Hoeh, Stewart and Guttman, 2002; Krebs, 2004). Furthermore, extensive phylogenetic analysis of M and F genomes might shed light on molecular evolutionary processes affecting these unique mtDNA genomes.

The genus *Pyganodon* (Unionidae: Unioninae) [previously included in the unionid genus *Anodonta*; see Hoeh, 1990] contains several conchologically similar species whose evolutionary relationships and distributions in eastern North America remain uncertain. Clarke (1981) recognized four species in this genus distributed across Canada, namely *P. cataracta* (Say, 1817), *P. fragilis* (Lamarck, 1819), *P. grandis* (Say, 1829) and *P. simpsoniana* (Lea, 1861). In a revised classification of the Eastern North American *Anodonta*, Hoeh (1990) indicated that *Pyganodon* includes two other species, *P. lacustris* (Lea, 1857) and *P. gibbosa* (Say, 1824) without mentioning *P. simpsoniana* [which could be conspecific with *P. grandis* (Clarke, 1981)]. In a recent survey of the genus

based on morphological characters and DNA sequences from nuclear (ITS1 and 2), F mitochondrial and M mitochondrial loci (cox1 and 16S) from southern Quebec, Cyr *et al.* (2007) (i) confirmed the distinction between *P. fragilis* and *P. cataracta*, (ii) confirmed the presence of *P. grandis* but not *P. simpsoniana* and (iii) found two additional *Pyganodon* mt lineages that they were unable to identify.

The aims of the present study are (1) to clarify the systematic relationships among *Pyganodon* lineages in northeastern North America using F and M mtDNA sequences from cytochrome c oxidase subunits I and II (cox1 and cox2); (2) to compare the phylogenetic signal obtained from M and F mitochondrial genomes in this group and (3) to place the pattern of genetic variation within and among lineages in a broad geographical and historical context.

4.3 Materials and methods

4.3.1 Populations sampled

Males and females, representing four *Pyganodon* species, were manually collected in northeastern North America from different populations of Newfoundland and Labrador, Quebec, Michigan, Minnesota, Wisconsin, New Jersey, Indiana, Maine, Maryland, Massachusetts and Pennsylvania (Table 4.1, Figure 4.1). Mussels were quickly transported to the laboratory and dissected. Tissues were frozen using liquid nitrogen in the field or whole mussels were preserved in 95–100% non-denatured ethanol in the field. Microscopic examination of gonadal tissue allowed for rapid determination of mussel gender (by the presence of eggs or sperm/sperm morulae). Dissected tissues samples were stored either at -80°C or in 95–100% ethanol at -20°C .

Table 4.1. Description and geographical coordinates of sampling sites for *Pyganodon lacustris*, *P. cataracta*, *P. fragilis* and *P. grandis* and numbers of individual per morphospecies sampled (see Figure 4.1 and Figures 4.9 to 4.12) documented in each site and morphospecies. L.: Lake, R.: River, P.: Pond, Res.: Reservoir, Co.: County, Gr: *P. grandis*, Fr: *P. fragilis*, Ca: *P. cataracta*, La: *P. lacustris*.

No.	Exact Site	Region	State/Province (Country)	Lat.	Long.	Species (N)	
						Female	Male
1	Atikamachus L.	Nord du Québec	Quebec (Canada)	53.5044	-77.7186	Gr (8)	Gr (9)
2	Matapédia L.	Bas-Saint-Laurent	Quebec (Canada)	48.5588	-67.5855	Fr (4)	Fr (11)
3	Macpès L.	Bas-Saint-Laurent	Quebec (Canada)	48.3119	-68.4869	Fr (9)	Fr (9)
4	Bellevue L.	Capitale-Nationale	Quebec (Canada)	47.1833	-72.2333		Ca (5)
5	Sainte-Marie L.	Outaouais	Quebec (Canada)	45.9561	-75.9305	Gr (2)	Gr (3)
6	Elgin L.	Estrie	Quebec (Canada)	45.7483	-71.3361	-	Fr (1)
7	Burin	Eastern	Nfld (Canada)	47.0504	-55.1632	Fr (1)	Fr (1)
8	Elephant L.	St. Louis Co.	Minnesota (USA)	48.1932	-92.7446	La (4)	La (6)

9	Black Duke L.	St. Louis Co.	Minnesota (USA)	48.2034	-92.8193	La (11)	La (6)
10	Pelican L.	St. Louis Co.	Minnesota (USA)	48.0595	-92.9098	La (9) Gr (3)	La (5) Gr (4)
11	Deer L.	Itasca Co.	Minnesota (USA)	47.3807	-93.6635	La (7) Gr (8)	La(12)Gr (2)
12	Pfeiffer L.	St. Louis Co.	Minnesota (USA)	47.7485	-92.4789	La (4) Gr (5)	La (1) Gr (1)
13	Prairie R.	Itasca Co.	Minnesota (USA)	47.2391	-93.4821	La (5) Gr (13)	La (9) Gr(6)
14	Rice R.	Aitkin Co.	Minnesota (USA)	45.7219	-94.2194	La (9) Gr (7)	La(26) Gr(6)
15	Twin lakes	Itasca Co.	Minnesota (USA)	45.2483	-94.2126	La (2)	-
16	Frontenac	Goodhue Co.	Minnesota (USA)	44.5269	-92.3297	Gr (4)	Gr (1)
17	Pogey L.	Mecosta Co.	Michigan (USA)	43.7962	-93.4821	Gr (8)	-
18	Mona L.	Muskegon Co.	Michigan (USA)	43.1767	-86.2602	Gr (2)	Gr (1)
19	Cedar R.	Gladwin Co.	Michigan (USA)	43.9932	-84.3704	Gr (4)	Gr (1)
20	Chetak L.	Sawyer Co	Wisconsin (USA)	45.7034	-91.4948	La (1) Gr (3)	La (1) Gr (1)
21	Kentuck L.	Vilas Co.	Wisconsin (USA)	45.9845	-89.0003	La (6)	La (4)
22	Annabelle L.	Vilas Co.	Wisconsin (USA)	46.2154	-89.6843	La (6) Gr (2)	La (2)
23	StGermain R.	Vilas Co	Wisconsin (USA)	45.9147	-89.5275	La (1) Gr (6)	La (5) Gr (6)
24	Mendota L.	Dane Co.	Wisconsin (USA)	43.1145	-89.4257	Gr (3)	Gr (1)
25	Ntwin	Vilas Co.	Wisconsin (USA)	44.2338	-89.1408	La (4) Gr (6)	La (4) Gr (4)
26	Pickering Creek	Chester Co.	Pennsylvania (USA)	40.1034	-75.5337	Ca (6) Fr (1)	Fr (3)
27	Prairie L.	Barry Co MI	Michigan (USA)	41.8586	-85.4037	Gr (5)	-
28	Lyons L.	Kalamazoo Co	Michigan (USA)	42.1884	-84.9668	Gr (8)	Gr (5)
29	Tilson farm P.	Hampshire Co	Maine (USA)	42.3903	-72.5327	Ca (3)	Ca (3)
30	Goose Bay	Labrador	Nfld. (Canada)	53.2945	-60.3618	Ca (1) Fr (5)	Fr (5)
31	Birds pond	Whittbourne	Nfld. (Canada)	47.4162	-53.5281	Fr (8)	Fr (4)
32	Indian L.	Washington Co.	Maine (USA)	45.4174	-69.2979	Ca (2)	Fr (1)
33	Bowstring R.	Itasca Co.	Minnesota (USA)	47.4972	-93.7435	La (4)	-
34	Beach pool	Emmet Co.	Michigan (USA)	45.7372	-84.8019	La (5)	-
35	Mantua Creek	Gloucester Co.	New Jersey (USA)	39.7088	-75.0923	Ca (6)	-
36	Loch Raven Res.	Baltimore Co.	Maryland (USA)	39.4177	-76.5383	Ca (1)	-
37	Fish L.	Calhoun Co.	Michigan (USA)	42.0494	-85.8600	Gr (2)	Gr (3)
38	Big Fork	Itasca Co.	Minnesota (USA)	47.8004	-93.5724	Gr (1)	-
39	Mill Creek	Jackson Co	Indiana (USA)	41.5664	-86.5490	Gr (3)	-

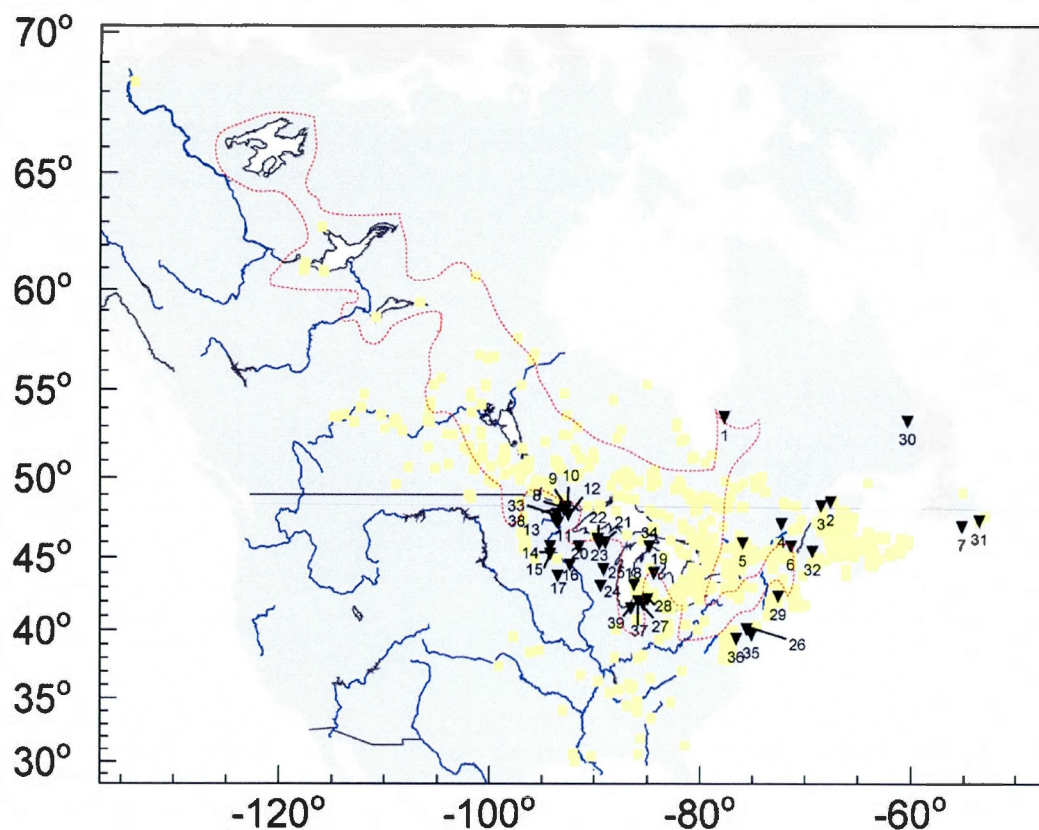


Figure 4.1 Sampling sites of *Pyganodon* spp. (black ▼) and geographical occurrence of *Pyganodon* spp. (yellow ■) based on Global biodiversity information facility (GBIF; Biodiversity occurrence data provided by: Canadian Museum of Nature, EMAN Provider, New Brunswick Museum, National Museum of Natural History and University of Kansas Biodiversity Research Center (Accessed through GBIF Data Portal, data.gbif.org, 2010-03-26). Sites are numbered as in Table 4.1. The maximum extent of the proglacial lakes (dashed line) is outlined.

4.3.2 DNA Extraction, amplification and sequencing

Total genomic DNA extractions were performed on the dissected tissues using a QIAGEN DNEasy Extraction Kit (QIAGEN Inc., Mississauga, Canada) following the manufacturer's protocol. Extracted total DNA quality and quantity were examined using 0.7% agarose gel electrophoresis and the total DNAs were stored at -20°C . Four distinct gene regions were analyzed in the present study: Fcox1, Mcox1, Fcox2, and Mcox2e. The complete F cytochrome c oxidase subunit II (Fcox2), Mcox2 and partial F and M cytochrome c oxidase subunit I (cox1) gene regions were amplified and directly sequenced without cloning. The following primer pair was used to amplify cox1: LCO22me2/HCO700dy2 (Walker *et al.*, 2006). These primers are complementary to very conserved regions of animal mtDNA (Folmer *et al.*, 1994) and have been shown to reliably PCR amplify mtDNA from unionoid bivalves (Hoeh *et al.*, 1996a; Hoeh *et al.*, 1998; Bogan and Hoeh, 2000; Hoeh, Bogan and Heard, 2001; Hoeh, Stewart and Guttman, 2002) and they facilitated the production of both Fcox1 (from ovarian or mantle tissue-based total DNAs) and Mcox1 (from testicular tissue-based total DNAs) amplicons. The primer pair UNIOND3 (5'-AGH SCK TTT GAR TGY GGK TTT GA-3') and FCOIPygR (5'-TGC CAR TAA CAA RTA YAA GTA-3') was used to amplify the Fcox2 gene from ovarian or mantle tissue-based total DNAs. Additionally, a third primer pair (MCOIIh 35F 5'-TTT ATR CCT RTK KTG TGT RGA RGC TGT-3'/PygMcox2eR 5'-TAY AAT CTT YCA ATR TCY TTA TGA TT-3') was used to amplify the hypervariable extension region of Mcox2 (= Mcox2e) [e.g., Walker *et al.*, 2006; Chapman *et al.*, 2008] from testicular tissue-based total DNAs. Some primers had a 5' M13 tail added for sequencing purposes. All amplifications were performed in 50- μl volumes of a solution containing 1X Qiagen buffer, 2.5 mM MgCl_2 , 200 μM each dNTP, 0.5 μM each primer, 2.5

U Qiagen Taq (QIAGEN Inc., Mississauga, Canada) and ~ 1 µg of template DNA. Thermal cycling conditions for *cox1* and *cox2* amplification were as follows: 94 °C for 3 min, followed by 35 cycles of 94 °C for 1 min, 54 °C for 1 min and 72 °C for 90 sec and a final extension at 72°C for 5 min. The resulting PCR products were gel purified using Qiagen QIAquick Gel Extraction Kit (Qiagen, Valencia, California, USA) or Promega Wizard PCR Kits (Promega, Madison, Wisconsin, USA) and sequenced at the Sequencing Platform (3730XL, Applied Biosystems) of McGill University (Québec, Canada; using the PCR primers) or at Geneway Research (Hayward, CA; using M13 primers).

4.3.3 Phylogenetic reconstruction

All sequences were were trimmed and assembled with the program BioEdit version 7.0.4.1 (Hall, 1999) and multiple sequence alignments were made using Clustal W (Larkin *et al.*, 2007) in MEGA 4.0 (Tamura *et al.*, 2007), using default parameters and Dialign version 2.2.1 (Subramanian, Kaufmann and Morgenstern, 2008) followed by manual alignment adjustments. Phylogenetic trees for the F and M genomes of *Pyganodon* were separately estimated using Bayesian inference (BI) and maximum likelihood (ML) analyses of concatenated nucleotide sequences from the *cox1* and *cox2* protein-coding genes. For the F sequence trees, sequences from 232 *Pyganodon* individuals and five outgroup sequences representing three *Anodonta* species (*Anodonta beringiana*, *A. woodiana* and *A. californiensis*) were using Clustal W in MEGA 4.0 and manually reviewed. In the same way, sequences from 184 *Pyganodon* individuals and those representing three outgroup species (*A. beringiana*, *A. woodiana* and *A. californiensis*) were aligned to produce the M sequence trees. The analyzed F-matrix had 1389

nucleotide positions (comprising 678 nucleotides from *cox2* and 711 nucleotides from *cox1*) and the M-matrix 1245 nucleotides positions (comprising 585 nucleotides from *cox2* and 660 nucleotides from *cox1*). These matrices are available from WRH.

Phylogenetic analyses were conducted using Bayesian inference (BI) via Mr. Bayes version 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Two independent simultaneous analyses were done using the GTR+G+I substitution model (Rodriguez *et al.*, 1990) for both the F and M sequence matrices. Searches were conducted for five million generations with six search chains each. The molecular data were partitioned by gene region and by codon position (= two gene regions, each with three codon positions for the *cox1* and *cox2* partitions) yielding a total of six partitions, and saving a total of 100,000 trees (one tree saved every 100 generations in each of the two analyses) from both the F and M sequence analyses. To allow each partition to have its own set of parameter estimates, *revmat*, *tratio*, *statefreq*, *shape*, and *pinvar* were all unlinked during the analyses. The BI analyses were terminated when the average standard deviation of the split frequencies fell below 0.020 and 0.011 for the F and M analyses, respectively. The 60,000 post-burnin trees (determined by examination of the log probability of observing the data x generation plots) were used to generate the F and M 50% majority-rule consensus trees. To allow each gene to evolve independently at its own rate, the option *prset ratepr=variable* was employed as per the recommendations of Marshall, Simon and Buckley (2006). Best maximum likelihood (ML) F and M trees were generated using GARLI version 0.951 (Zwickl, 2006) with no data partitioning using default settings except for the following: autoterminate run 1,000,000 generations post last improved topology, lnL increase for

significantly better topology = 0.0001 and score improvement threshold = 0.0005. Garli was also used to generate a 200-replicate ML majority-rule bootstrap (Felsenstein 1985) trees using default settings except for the following: InL increase for significantly better topology = 0.001 and score improvement threshold = 0.005, based on analyses of the F and M sequence matrices with no data partitioning.

4.3.4 Sequence-based species delimitation

We implemented two different empirical methods to delimit species, including a general mixed Yule-coalescent method (GMYC) and a statistical parsimony network analysis. The general mixed Yule coalescent model (Pons *et al.*, 2006; Fontaneto *et al.*, 2007) is a statistical procedure for estimating species boundaries directly from phylogenetic trees. Equations describing two types of lineage evolution, speciation (a stochastic birth-only or Yule coalescent model) and a model that described population (a neutral coalescent model; microevolution) are combined and applied on an ultrametric tree. The procedure identifies a threshold value for the shift in branching rate from macroevolution to microevolution. The number of shifts and their location on the phylogenetic tree define the species boundary. A standard log-likelihood ratio test is then used to assess whether the stepped model provides a significantly better fit than the null model of no such shift in branching process. The analysis was carried out using R package SPLITS (Species Limits by Threshold Statistics) available at <http://r-forge.r-project.org/projects/splits/>. To build the clock-constrained tree, we used the Bayesian consensus tree as described before. From this tree, identical haplotype and out-group sequences were removed and converted to an

ultrametric tree using penalized likelihood in r8s 1.7 (Sanderson, 2003), with the optimal smoothing parameter determined by cross validation between 10^{-5} and 10^5 .

For the second algorithmic method of sequence-based species delineation, patterns of sequence variation were analyzed for the presence of species-level groups by identifying independent networks using parsimony (Templeton, Crandall and Sing, 1992) as implemented in the software TCS version 1.21 (Clement, Posada and Crandall, 2000). This procedure partitions the data into independent networks of closely related haplotypes connected by changes with a $< 95\%$ probability to be non homoplasmic. Alignment gaps were treated as missing data.

We measured divergence (uncorrected p-distance) within and among lineages using MEGA version 4. To test whether genetic differentiation among putative species was significant, we used analysis of molecular variance (AMOVA; Excoffier, Smouse and Quattro, 1992) implemented in ARLEQUIN version 3.01 (Excoffier, Laval and Schneider, 2005). For combined *cox1* and *cox2* sequence from each *Pyganodon* lineage, standard measures of polymorphic sites, nucleotide diversity and sequence diversity were obtained from the PGEToolbox for MATLAB (Cai, 2008) and ARLEQUIN version 3.11 (Excoffier, Laval and Schneider, 2005).

4.4 Results

In this study, DNA samples for the *cox1* gene fragment within the genus *Pyganodon* were analyzed to produce 224 F-haplotypes and 178 M haplotypes. We also obtained 107 F haplotypes and 164 M haplotypes of a fragment of the *cox 2* gene. Four morphological species (hereafter called morphospecies) were identified and designated as *P. cataracta*, *P. fragilis*, *P. lacustris* and *P. grandis* (Figure 4.2). The molecular divergence between morphospecies is presented in Table 4.2 for F haplotype and in Table 4. 3 for M haplotype. The further subdivisions within morphospecies are based on phylogenetics analysis.

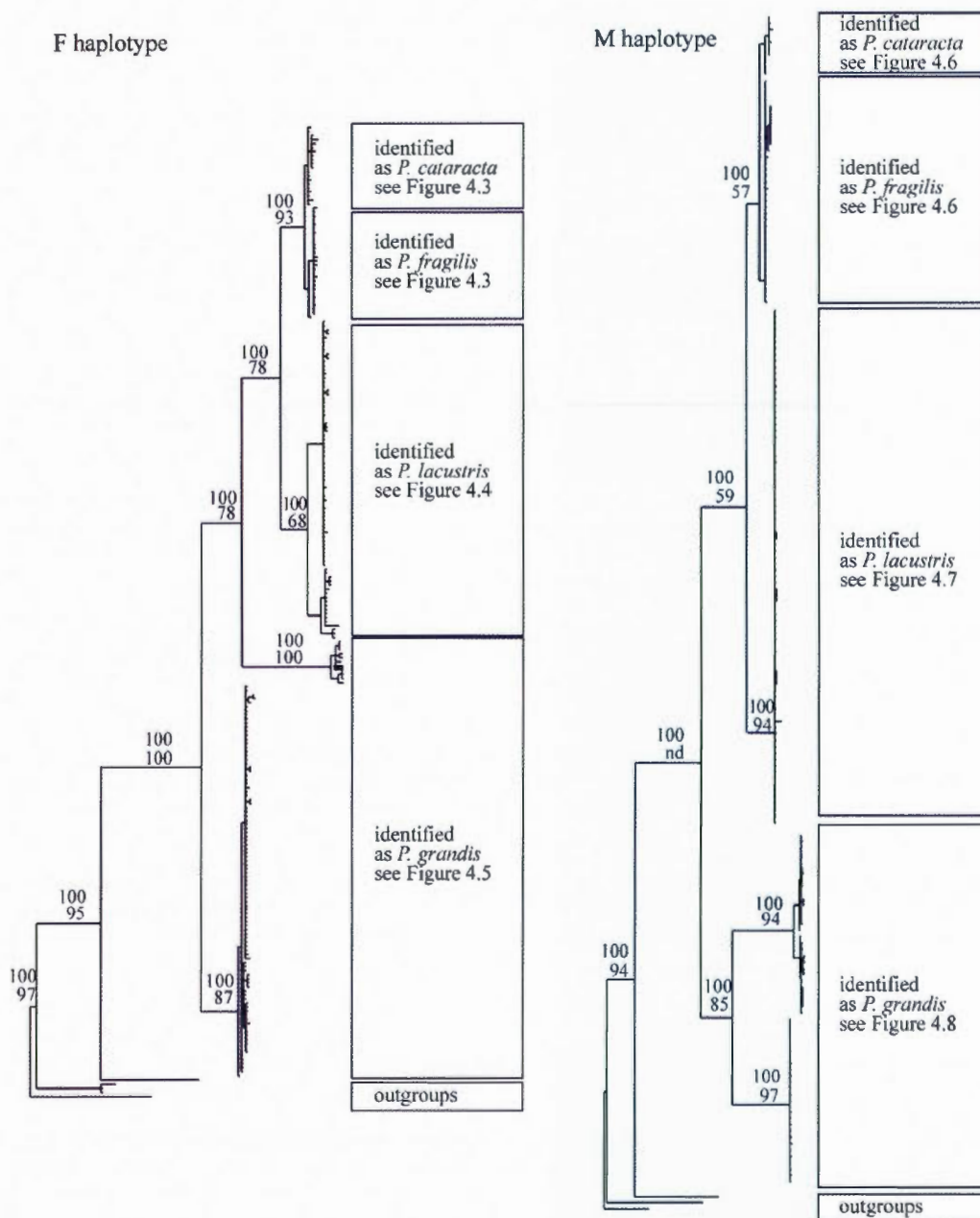


Figure 4.2. Bayesian tree of the F and M *Pyganodon* complex inferred from *cox1* and *cox2*. ML bootstrap values (below) and BI posterior probabilities (x100, above) are indicated at the branches. Boxes indicate to which species the specimens were identified morphologically prior to the genetic analysis.

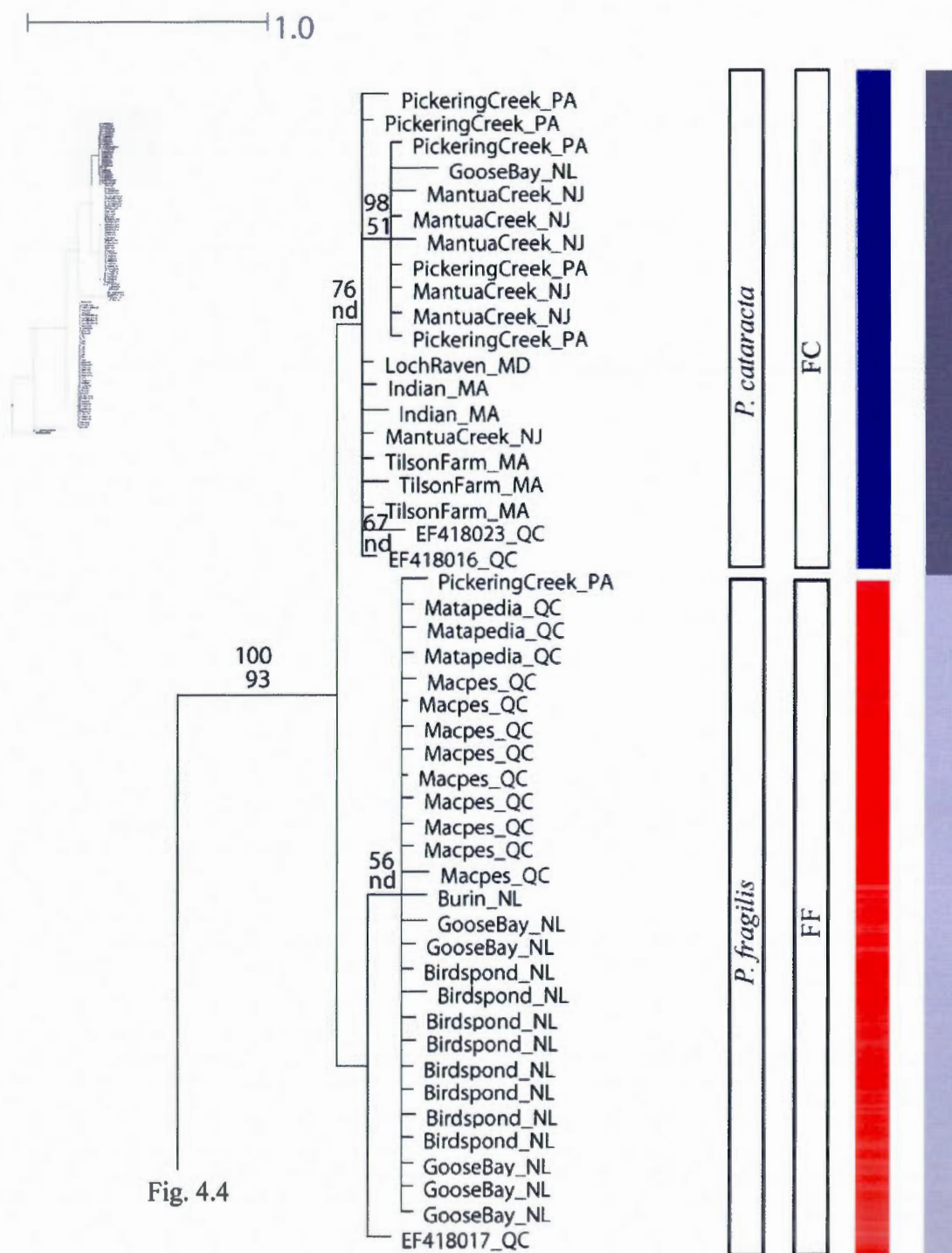


Fig. 4.4

Figure 4.3 Bayesian majority-rule consensus tree of the *F. Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represent the morphospecies, the second represent the genome (F vs. M). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. In the last column, two tones of gray represent a connected network divided in two distinct subnetworks.

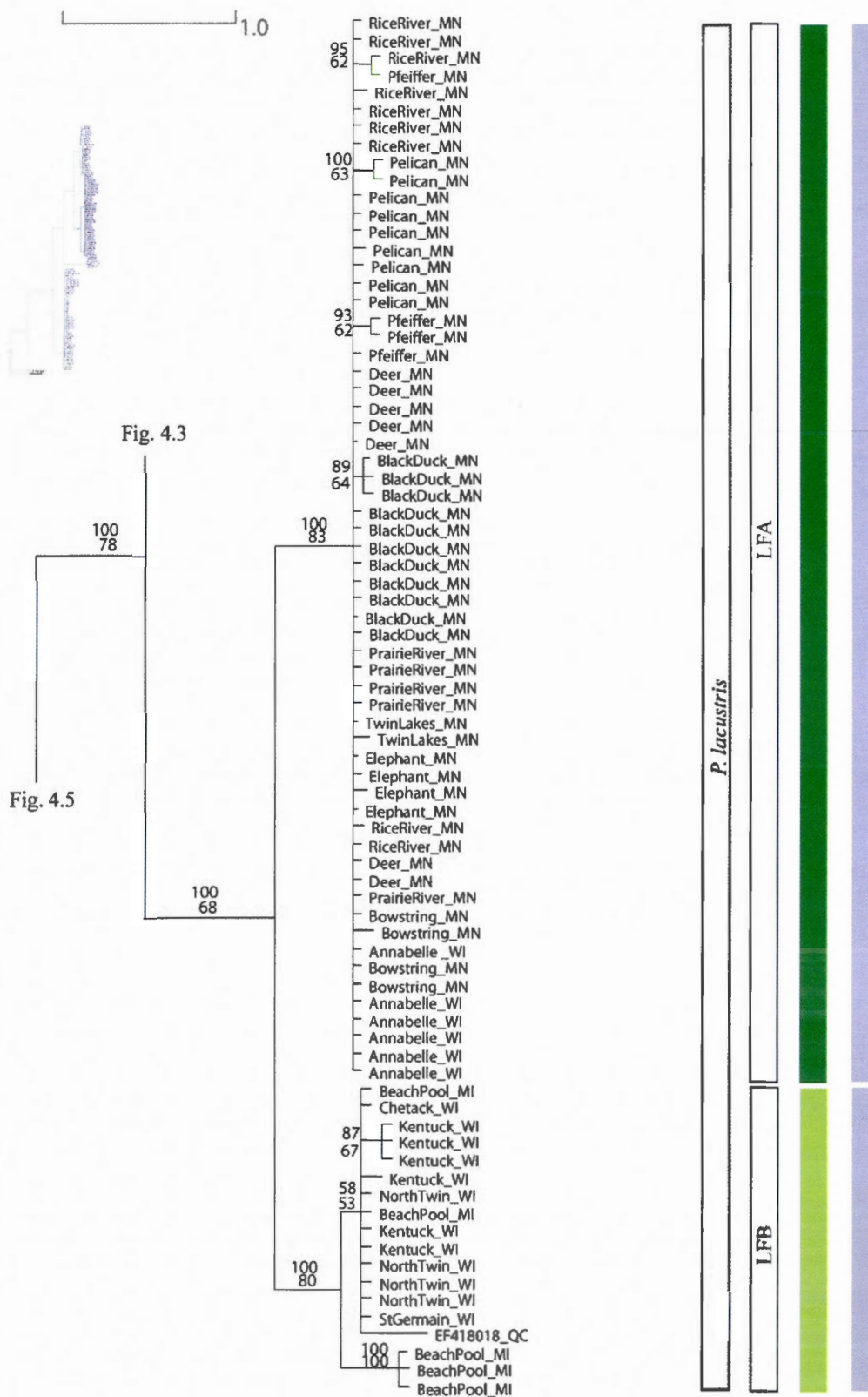


Figure 4.4 Bayesian majority-rule consensus tree of the F *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represent the morphospecies, the second represent the genome (F vs. M) and the third represents a subclade (A or B). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks.

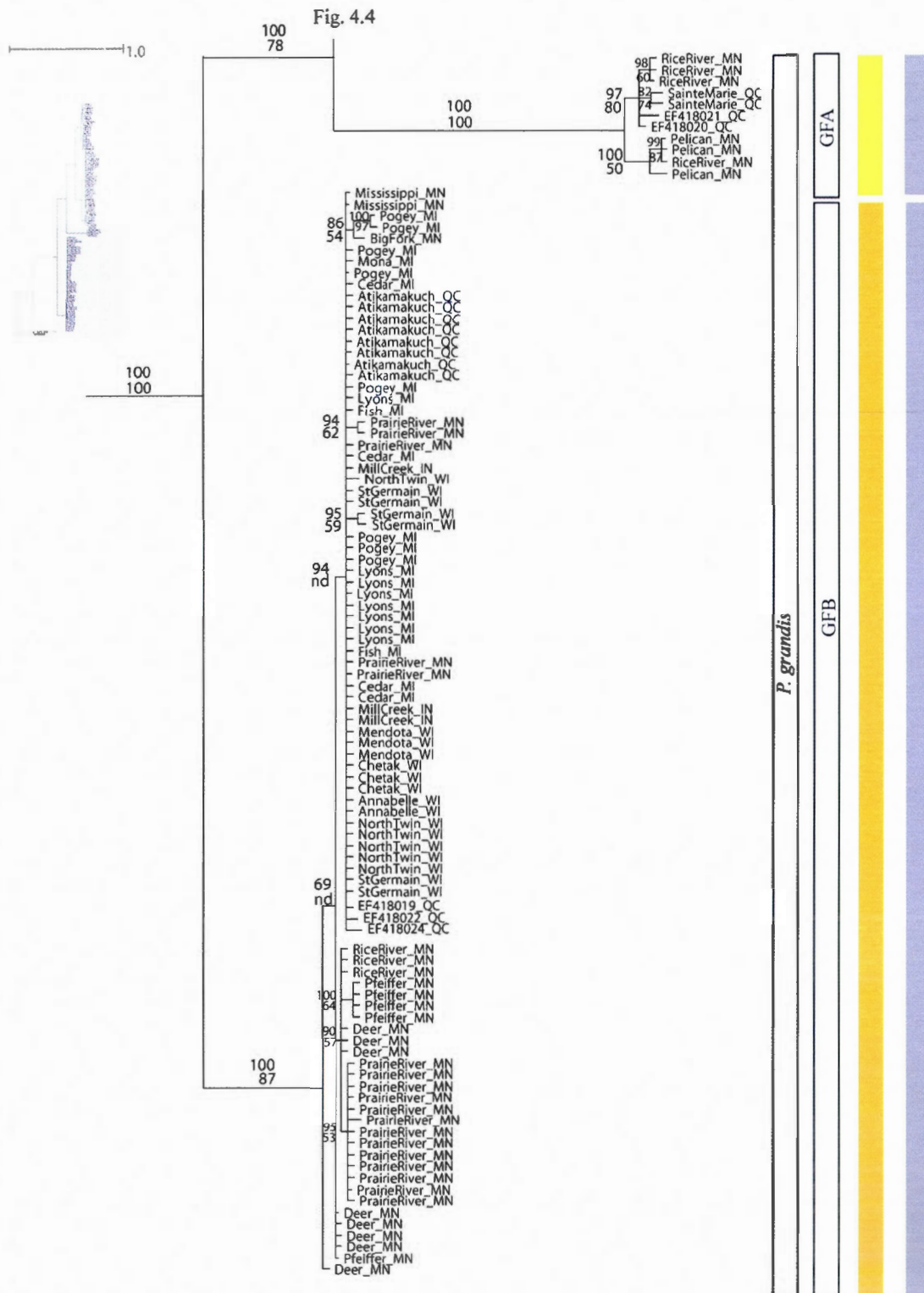


Figure 4.5 Bayesian majority-rule consensus tree of the F *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M) and the third represent a subclade (A, B or C). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks.

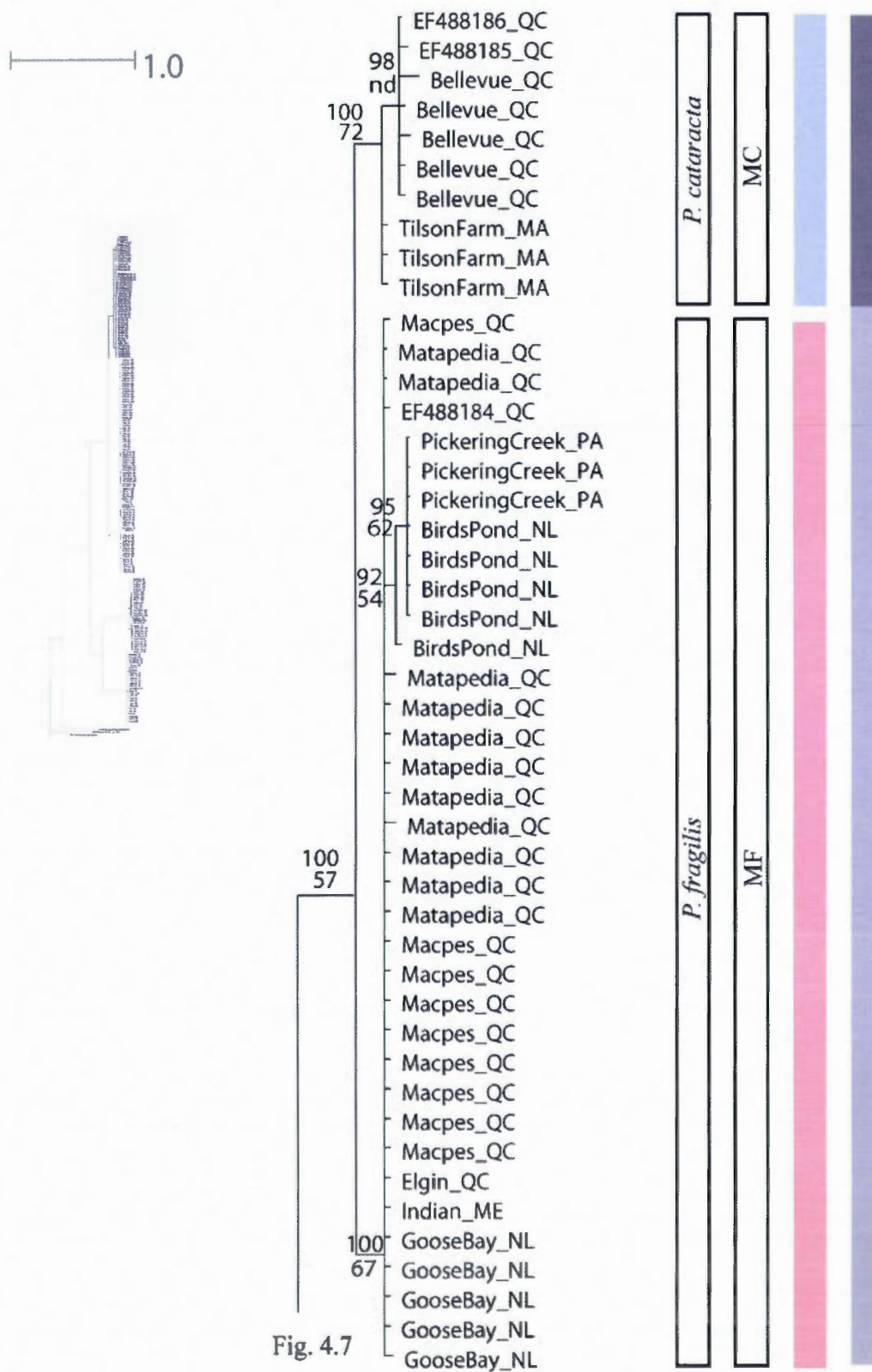


Figure 4.6 Bayesian majority-rule consensus tree of the M *Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. In the last column, two tones of gray represent a connected network divided in two distinct subnetworks.



Fig. 4.8

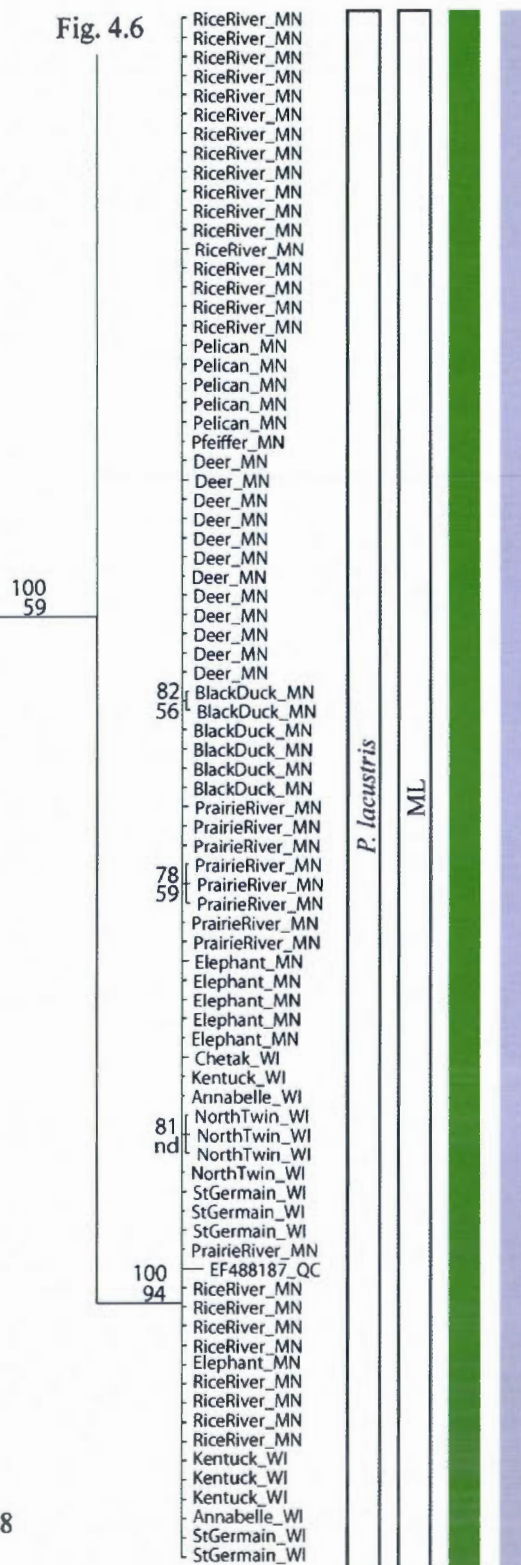


Fig. 4.6

Figure 4.7 Bayesian majority-rule consensus tree of the *M. Pyganodon* complex. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. GenBank accession numbers come from Cyr *et al.* (2007). The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represent the genome (F vs. M). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks.

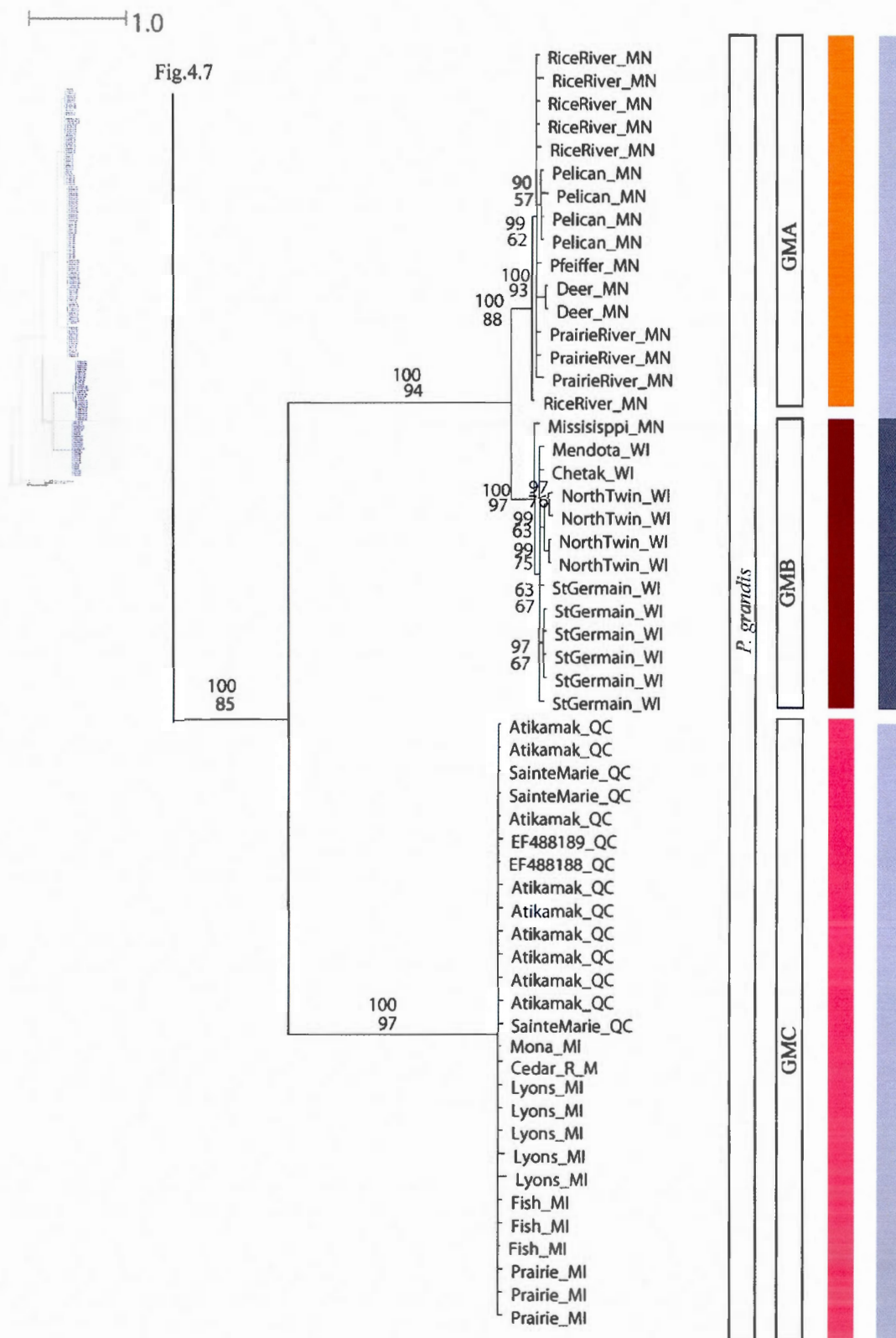


Figure 4.8 Bayesian majority-rule consensus tree of the M *Pyganodon* complex inferred from *cox1* and *cox2*. ML bootstrap values (below) and BI posterior probabilities (x100, above) are indicated at the branches. Terminal labels indicate collection site and abbreviation of US states or Canadian provinces. The first column to the right of the tree indicates species as identified by morphology. The second column refers to the name of clades (the first letter of the acronym represents the morphospecies, the second represents the genome (F vs. M) and the third (if necessary) represents a subclade (A, B or C). The third column shows the boundaries identified by the likelihood analysis (GMYC method). The fourth column indicates membership in statistical parsimony networks. In the last column, two tones of gray represent a connected network divided in two distinct subnetworks.

4.4.1 Phylogenetic analysis

The BI and ML phylogenetic analyses produced similar topologies for F and M data sets. Therefore, in the following discussion, we use as our point of reference the majority-rule consensus phylogenetic trees obtained by Bayesian inference (Figure 4.2; for a detailed structure of the F and M trees, see Figures 4.3–4.8).

In the F tree, two relatively differentiated clades (hereafter LFA and LFB; see Figure 4.4) in *P. lacustris* morphospecies (Table 4.2 and Figure 4.3). The branching patterns in the M tree were different where *P. lacustris* was comprised of a single clade containing very similar haplotypes. However, in the M tree, the *P. lacustris* clade is the sister group to the *P. cataracta* + *P. fragilis* lineage as in the F tree. Furthermore, in the F tree, two, relatively differentiated clades (hereafter named GFA and GFB) were found in *P. grandis* rendering the species paraphyletic (Figures 4.2 and 4.5; Table 4.2). The presence of two, non-sister clades in the F mitochondrial genomes obtained from *P. grandis* morphospecies indicates cryptic mitochondrial

lineages. Specifically, the *P. grandis* clade GFA is the sister group to the F clade containing *P. cataracta* + *P. fragilis* + *P. lacustris* morphospecies. The larger *P. grandis* F clade (GFB) is the sister group to the monophyletic group containing all other *Pyganodon* F sequences. However, in the M tree (Figures 4.2 and 4.8), *P. grandis* is monophyletic but the lineage comprises three clades (hereafter GMA, GMB and GMC). Divergence among those three clades ranges from 10.13 to 12.51 % (Table 4.3). The *P. grandis* A clade is the sister group to the monophyletic group containing all other *Pyganodon* M sequences.

Table 4.2. Morphospecies and divergence in percentage (%) between F.

	<i>P. cataracta</i>	<i>P. fragilis</i>	<i>P. lacustris</i> LFA	<i>P. lacustris</i> LFB	<i>P. grandis</i> GFA
<i>P. fragilis</i>	0.43				
<i>P. lacustris</i> LFA	4.46	4.10			
<i>P. lacustris</i> LFB	5.05	4.66	1.64		
<i>P. grandis</i> GFA	8.38	7.99	10.28	11.01	
<i>P. grandis</i> GFB	9.58	9.43	9.16	10.69	9.60

Table 4.3. Species divergence in percentage (%) between M.

	<i>P. cataracta</i>	<i>P. fragilis</i>	<i>P. lacustris</i>	<i>P. grandis</i> GMA	<i>P. grandis</i> GMB
<i>P. fragilis</i>	1.81				
<i>P. lacustris</i>	5.20	4.71			
<i>P. grandis</i> GMA	11.25	10.30	11.58		
<i>P. grandis</i> GMB	10.93	9.99	11.41	12.51	
<i>P. grandis</i> GMC	12.54	12.63	12.82	10.13	10.52

The two other morphospecies, *P. cataracta* and *P. fragilis*, are reciprocally monophyletic and the *P. cataracta* + *P. fragilis* clade is the sister group to *P. lacustris* in both the F and M trees (Figures 4.2). However, there is relatively little genetic divergence between the *P. cataracta* and *P. fragilis* F and M sequences (Tables 4.2 and 4.3) such that nodal support values are robust for species' monophyly only in the M tree.

4.4.2 Sequence-based species delimitation

Analyses of branch length dynamics were also applied to better delimit putative species. In the lineage-through-time plots, the shift in branching rates is clear in both M and F analyses. Furthermore, for both analyses, the GMYC model of lineage evolution was statistically a better fit to the datasets than the null model of uniform coalescent branching. Likelihood values for the F analysis were $\log L = -25.362$ for the GMYC model and $\log L = -63.228$ for the null model; $2\Delta\log L = 75.732$, χ^2 test, d.f. = 3, $p \ll 0.001$. For the M analysis, likelihood values were $\log L = -7.259$ for the GMYC model and $\log L = -55.281$ for the null model; $2\Delta\log L = 96.043$, χ^2 test, d.f. = 3, $p \ll 0.001$. For the F sequences, the analysis revealed a total of six species (interval for the number of clusters 6 to 8) and for the M sequences, a total of six species with a confidence interval of six.

The clusters identified by these shifts closely match with those identified by our second algorithmic sequence-based method for species delineation, namely the network analysis (see the correspondence between columns 3 and 4 in Figures 4.3–4.8). Statistical parsimony network construction within the 95% connection limit conducted separately for F and M sequences resulted in five and four unconnected networks respectively

(Figures 4.3–4.8). Both F and M haplotype networks regroup *P. cataracta* and *P. fragilis* morphospecies into one independent network formed by two connected subnetworks. Similarly, while the GMYC method based on M sequences resolve *P. grandis* morphospecies into three putative species, the parsimony network analysis revealed two unconnected networks but one of them is actually constituted of two subnetworks that correspond to GMA and GMB clades on the tree.

4.4.3 Biogeography

P. cataracta (Figure 4.9: clades FC and MC) and *P. fragilis* (Figure 10: clades FF and MF) occupy the eastern part of the area of study while the two other species are sympatric, mostly around all the Great Lakes (Figures 4.11 and 4.12). There are two localities (Goose Bay, Labrador and Pickering Creek, Pennsylvania) where some female individuals clustered in *P. fragilis* and some others clustered in *P. cataracta* putative species based on mtDNA sequences but were indistinguishable according to their morphologies. *P. lacustris* and *P. grandis* were often found in same locations in our survey. The range of *P. grandis* (clades GFB and GMC) extends north to the James Bay region while we found *P. cataracta* (clade FC) and *P. fragilis* (clades FF and MF) in the Labrador region. The F genome of the *P. lacustris* morphospecies displays two distinct mitochondrial lineages (clades LFA and LFB) and the geographical distributions of these two F clades show a clear structure where the Wisconsin, Michigan and Quebec *P. lacustris* populations (LFB) are differentiated from those in Wisconsin and Minnesota (LFA) while the homogeneous M genome is widespread over the distribution range of the morphospecies (see Figure 4.11). One *P. grandis* F genome clade, GFB, is widely distributed all around the Great Lakes region and is also found in the

James Bay region (Figure 4.12). The second *P. grandis* F genome clade, GFA, was less common and was detected in specimens from the Rice and Pelican River drainages in Minnesota as well as in the southwestern part of the St-Lawrence drainage basin (Figure 4.12). The three clades from analyses of *P. grandis* M genomes clearly depict a geographic structure. The GMC clade is broadly distributed to the east of Lake Michigan while GMA and GMB are principally found to the west of Lake Michigan (Figure 4.12). The GMB clade is spatially clustered principally in Wisconsin while GMA is found in Minnesota (Figure 4.12).

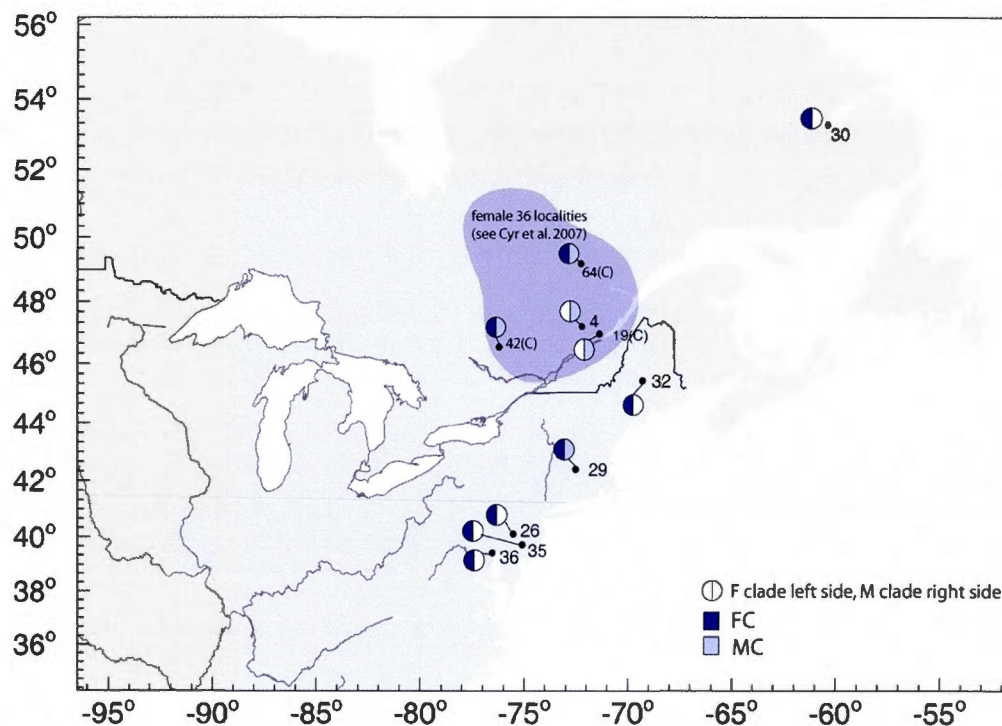


Figure 4.9 Geographic distribution of *Pyganodon cataracta* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). Color codes and clade names follow Figures 4.3 and 4.6. An empty half circle indicates that no data exist for that location. Sites are numbered as in Table 4.1. Symbol (C) Data from Cyr *et al.* (2007). The purple zone indicates area of the 36 localities from Cyr *et al.* (2007).

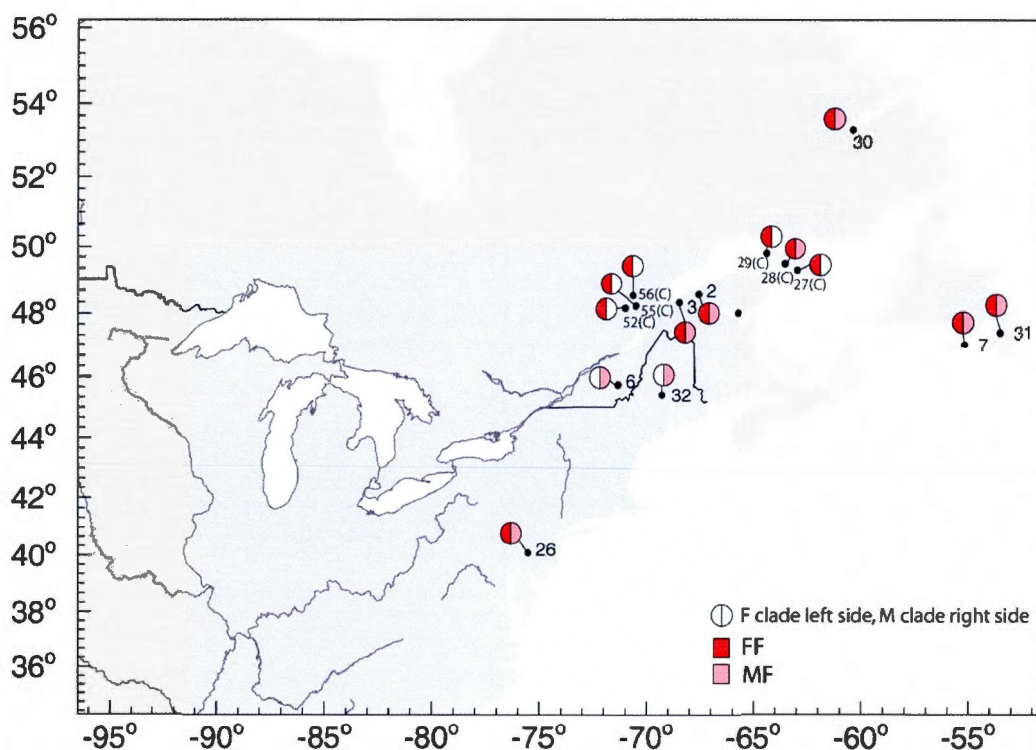


Figure 4.10 Geographic distribution of *Pyganodon fragilis* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). An empty half circle indicates that no data exist for that location. Color codes and clade names follow Figures 4.3 and 4.6. Sites are numbered as in Table 4.1. Symbol (C) means data from Cyr *et al.* (2007).

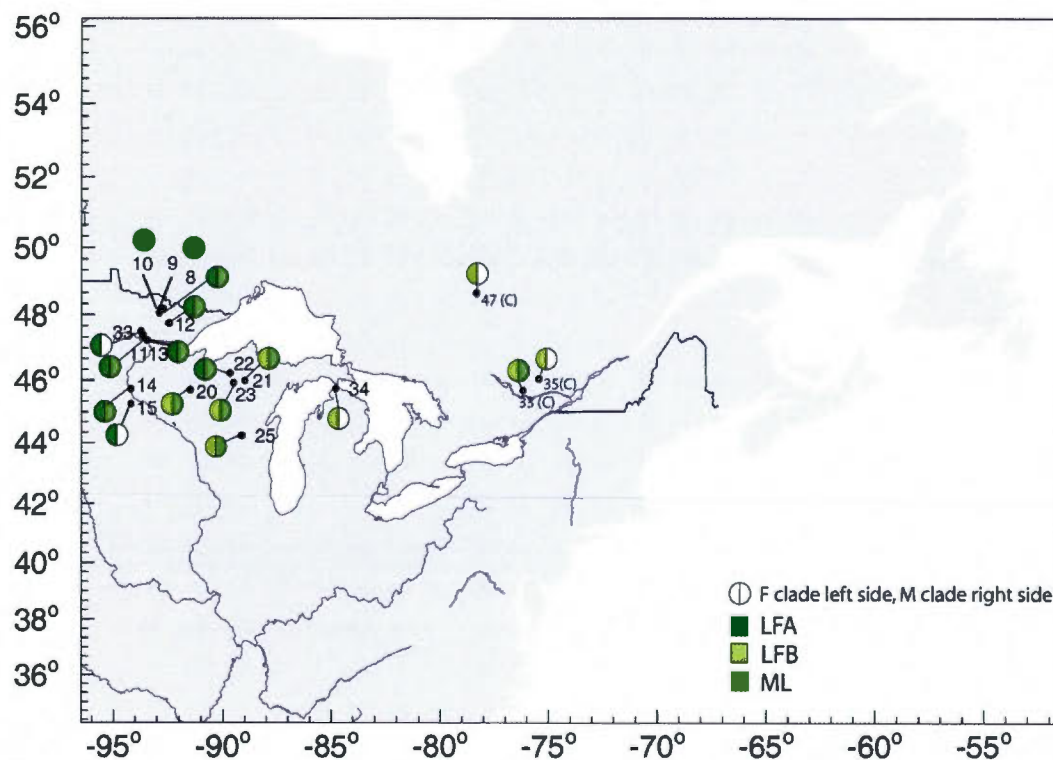


Figure 4.11 Geographic distribution of *Pyganodon lacustris* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). An empty half circle indicates that no data exist for that location. Color codes and clade names follow Figures 4.4 and 4.7. Sites are numbered as in Table 3.1. Symbol (C) means data from Cyr *et al.* (2007).

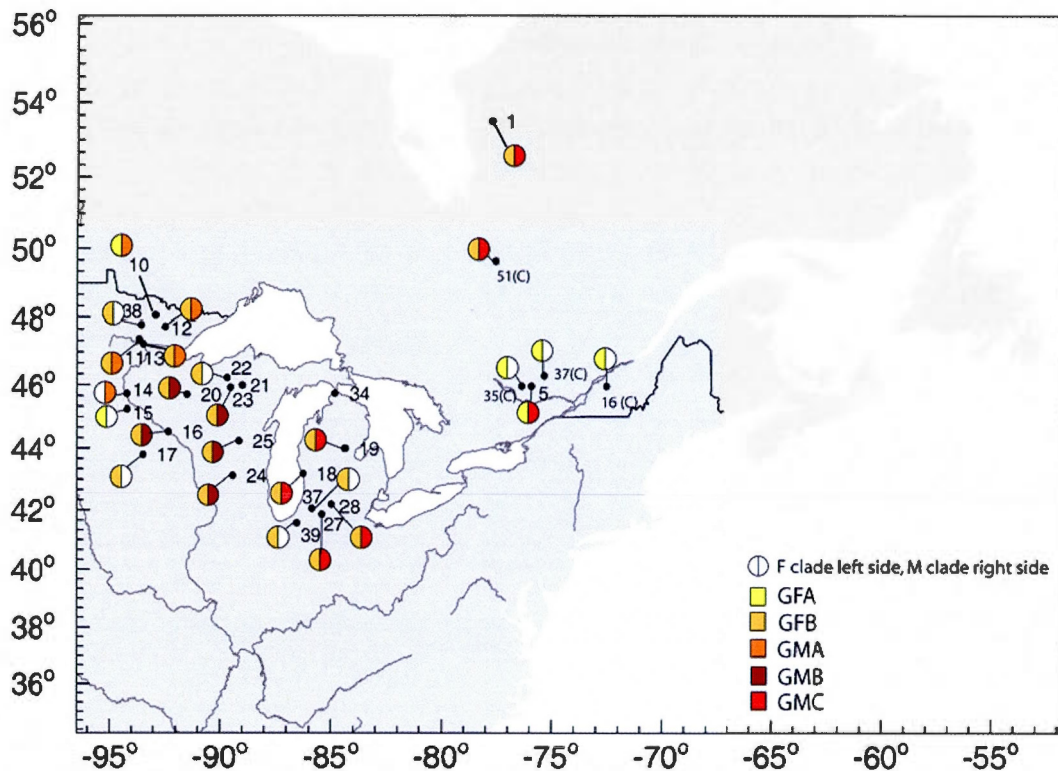


Figure 4.12 Geographic distribution of *Pyganodon grandis* morphospecies clades inferred from F (left side of the symbol) and M genomes (right side of the symbol). An empty half circle indicates that no data exist for that location. Color codes and clade names follow Figures 4.5 and 4.8. Sites are numbered as in Table 3.1. Symbol (C) means data from Cyr *et al.* (2007).

4.4.4 Variation in F and M genome within and among clades

For the F genomes, most of the variation was found within populations (54.33%; Table 4.4). Variation among populations within clades accounted for 24.45% leaving 21.22% of the total variation among clades. For the M genomes, the partitioning of variation is different. As for the F genomes, most of the variation was found within populations (48.22%; Table 4.4) but there was little variations among populations within clades (6.92%). The variation among clades comprised the remaining 44.85%. All values were significant (AMOVA $p < 0.001$; 10,000 permutations).

Table 4.4. Hierarchical AMOVA for the M and F data sets from *Pyganodon* morphospecies/clades.

Source of variation	Male			Female		
	df	Var	%	df	Var	%
Among clades	5	119.40	44.85*	5	26.10	21.22
Among population within clades	37	18.43	6.92*	47	30.07	24.45
Within populations	136	128.38	48.22*	175	66.82	54.33

* $p < 0.001$.

Pyganodon cataracta (FC) and *P. fragilis* (FF) diverged from one another by only 0.43% (uncorrected p -distance) according to the F data (Table 4.2). The M genome was, however, nearly four times more divergent between these two morphospecies/clades than their F counterpart (Table 4.3). Genetic distance between the two F clades of *P. lacustris* (LFA x LFB) was 1.64% and they both diverged from *P. cataracta* and *P. fragilis* by ~ 4.50% and from *P. grandis* clades (GFA and GFB) by ~ 9 – 11%. Values for these comparisons are similar but slightly higher for the M genomes, i.e. ~ 4.9% and ~ 11 – 12% respectively. The genetic distances between *P. grandis* clades are the highest within a single morphospecies averaging 9.60% and 12.51% for F and M genomes, respectively. The patristic distances between the *Pyganodon grandis* clades and all other clades are the largest in our study.

4.5 Discussion

We studied maternally and paternally inherited mtDNA sequence variation in the *Pyganodon* complex from northeastern North America. While we identified four species based on morphology, namely *Pyganodon grandis*, *P. lacustris*, *P. fragilis* and *P. cataracta*, the phylogenetic analyses of F and M mitochondrial sequences revealed the presence of six divergent clades, which has forced us to reconsider the actual number of species. However, our data do not lead to a simple classification into six species as there is extensive disagreement between maternally and paternally inherited genes with respect to inter- and intraspecific relationships in the *P. lacustris* and *P. grandis* group.

Sequence-based species delineation is particularly useful for taxa in which phenotypic plasticity makes morphological identification difficult. DNA sequence-based taxonomy includes several quantitative methods such as population aggregation analysis [PAA; Davis and Nixon (1992)], cladistic haplotype aggregation (Brower, 1999) and the Wiens and Penkrot methods [WP; Wiens and Penkrot (2002)]. These procedures aggregate populations lacking discrete differences into a single species. In these methods, species are recognized based on fixed nucleotide differences unique to sets of populations, which are defined *a priori* based on morphological, geographical or ecological information (Sites and Marshall, 2004). Given the possibility of cryptic diversity (Cyr *et al.*, 2007), morphological similarities and possibly sympatric clades/morphospecies, defining sets of population *a priori* to fit into those procedures could introduce a strong bias in the species delineation. We, therefore, rely on quantitative methods that do not require prior assumptions of population boundaries. The GMYC and the Templeton

parsimony methods gave a total of six putative species from analyses of the F sequences and also a total of six putative species from analyses of the M sequences. However, the predicted number of distinct putative species within each of the recognized morphospecies is not congruent between M and F genomes.

4.5.1 Recent speciation of *P. cataracta* and *P. fragilis*

The degree of reproductive isolation between *P. fragilis* and *P. cataracta* has been subject of controversy (Clarke and Rick, 1963; Kat, 1983; Kat, 1986). They are morphologically very similar and difficult to distinguish based on shell characteristics (Hanlon and Smith, 1999; Cyr *et al.*, 2007). Because beak sculpture is subject to erosion, the number of beak loops overlaps (Clarke, 1981), periostracum color varies with habitat (Hanlon and Smith, 1999) and stomach anatomy does not reliably distinguish them (Hanlon and Smith, 1999), identification of these species based on these conchological and anatomical criteria remains questionable. However, both of our mitochondrial datasets distinguish two putative species with relatively robust ML and BI nodal support values present in the F tree (93% /1.00 and 57%/1.00 for the F and M tree respectively). In addition, the GMYC species delimitation analysis clearly indicates a specific distinction between those clades while the Templeton parsimony analysis tends toward the same conclusion. Therefore, our data support a distinction between *P. fragilis* and *P. cataracta*, but these species appear to have diverged relatively recently from their common ancestor.

In some drainages [i.e., Goose Bay drainage (Labrador) and Pickering Creek (Pennsylvania)], we could not distinguish distinct species based on morphology. However, mtDNA from different individuals of a population clustered in both the *P. cataracta* and *P. fragilis* mt clades (Figures 4.3 and 4.6) and multiple Pickering Creek *P. cataracta* individuals' M sequences were nearly identical to those from some *P. fragilis* individuals from Newfoundland (Figure 4.6). The puzzling distribution of haplotypes could be due to either (1) incomplete lineage sorting or (2) interspecific hybridization. The apparent recency of divergence between these two species makes either hypothesis plausible. However, the relatively large geographic distance between these populations and the fact that they are relatively distant from a potential hybrid zone for these two species, best support the former hypothesis.

As reported in Cyr *et al.* (2007), the M sequence divergence between these two *Pyganodon* species was considerably greater than that observed in their F sequences. In fact, the M sequences were more divergent than the F sequences in all interspecific comparisons of *Pyganodon*. However, the *P. cataracta*–*P. fragilis* M divergence/F divergence ratio is, by far, the largest in this study. Cyr *et al.* (2007) suggest that the divergence in the female genomes has been erased following a relatively ancient hybridization event. An alternative explanation for this pattern could be ongoing speciation where both species were isolated from each other in different refugia during the last glaciation period. Possibly, the M mitochondrial genome evolved relatively faster than the F genome during the early stages of speciation. Consistent with this hypothesis, the M genome could, at this time, be under a more relaxed selective constraint (Stewart *et al.*, 1996; Zouros, 2000) which would allow more mutations to accumulate.

4.5.2 Non-monophyly/cryptic speciation in *P. grandis*

Cyr *et al.* (2007) identified predominantly only one lineage for *P. grandis*. We documented two relatively distinct (~10% uncorrected sequence divergence) F lineages in the *P. grandis* morphospecies rendering this species paraphyletic and basal in the genus *Pyganodon*. One *P. grandis* F genome clade (GFA) was less common and was detected in specimens from two river drainages in northern Minnesota as well as in the southwestern part of the St. Lawrence River drainage basin in Quebec. Therefore, the geographic distribution of the GFA clade appears discontinuous. This could be accurate or, alternatively, GFA might be more widely distributed to the north of the Great Lakes across the Ontario region (Canada), an area not covered by our sampling. If this latter hypothesis is verified, the distribution of GFA could correspond to *P. simpsoniana*, a species supposedly closely related to *P. grandis* (Clarke, 1981). Despite its description in the literature, our data does not allow us to confirm the existence of this species. Haplotypes from one of the unidentified *Pyganodon* F lineages of Cyr *et al.* (2007) (therein called species lineage B) reside in our GFA clade. Our second *P. grandis* F clade, GFB, was the most widely distributed all around the Great Lakes region and was also found in the James Bay region. This situation, with two relatively divergent mt lineages within one species, is highly unusual as intraspecific divergence in animal mtDNA is usually 10 times lower than interspecies divergence (Hebert *et al.*, 2004). The level of F mtDNA sequence divergence within the morphospecies *P. grandis* is of the same magnitude as that observed between any F clades of *P. grandis* and *P. lacustris* and higher than that of all other interspecific F lineage comparisons in *Pyganodon*.

In the *P. grandis* M genomes, there are not two but three highly divergent clades (GMA, GMB and GMC) or three clades with two being relatively genetically similar (GMA and GMB) but, together, divergent from the third clade (GMC). The M genomes possess slightly greater levels of intraspecific variation than do the F genomes of this morphospecies. However, unlike the situation in the F tree (Figure 4.5), the three *P. grandis* M clades form a monophyletic group (Figure 4.8). A greater level of intraspecific variation in the M compared to the F genomes is in accordance with what has been previously demonstrated in other DUI systems [e.g. *Mytilus*: Stewart *et al.*, 1995; Sutherland *et al.*, 1998, *Tapes* (= *Venerupis*): Passamonti and Scali, 2001].

The presence of divergent F and M genomes in *P. grandis* is consistent with the hypothesis that interspecific hybridization has occurred and that mt genes from distinct species have introgressed into *P. grandis*. When regarding the geographical pattern of distribution, it appears that both F genomes could be geographically paired with different M genomes and vice versa. For example, GFA occurs with GMC in Quebec, but also occurs with GMA in Minnesota. These facts point more toward a complex of species that share both M and F genomes. It is surprising that such sharing could occur with a high level of divergence between GFA and GFB groups in F genomes and GMA, GMB and GMC groups in M genomes. Yet another possibility is that *P. grandis* originated via the fusion of previously divergent gene pools (for both mtDNA and nuDNA). The Cyr *et al.* (2007) nuclear sequence data (from ITS) indicate that their "species lineage B" and "*P. grandis*" F genomes are found associated with a singular nuclear background. This finding is consistent with the hypotheses that (1) *P. grandis* is a single species and (2) mtDNAs from other species have introgressed into *P. grandis*. The fact that

the branching patterns for the *P. grandis* F and M genomes are significantly different support this hypothesis as the five mtDNA lineages we see in *P. grandis* could have been the result of a comingling of gene pools from three species, e.g., the "original" *P. grandis* which contributed an F and an M genome, a *P. lacustris*-like species to account for the *P. grandis* F lineage that is the sister group to *P. lacustris*+*P. cataracta*+*P. fragilis* lineages and a *P. grandis*-like species (possibly *P. simpsoniana*) to account for one of the three *P. grandis* M lineages those are sister lineages. Introgressive hybridization could have been facilitated by the dramatic ecological disturbances caused by the filling and draining of the ancient proglacial Lake Ojibway and Lake Agassiz, which occurred toward the close of the Wisconsinan glacial episode (Barber *et al.*, 1999; Fisher and Smith, 1994; Veillette, 1994). The filling of these large basins would have facilitated contiguity between previously isolated river drainages thus allowing divergent *Pyganodon* populations to exchange genes and individuals to disperse widely across the considerable breadth of these ancient lakes. If the above scenario occurred, i.e., augmentation of gene pools via interspecific hybridization, it could explain the protean nature of *P. grandis*, possibly the most conchologically variable unionid species in North America.

4.5.3 Sweep in *P. lacustris* M genome

The *P. lacustris* F genome displays two geographically distinct clades while the homogeneous M genome has a widespread geographic distribution. Even though we cannot rule out the possibility that a divergent M genome was not sampled, it appears that the M and F genomes of *P. lacustris* have undertaken different evolutionary trajectories where the F genomes could have been isolated from each other long enough to diverge. A single M

lineage appears to have recently invaded the M genome space in this species inside the territory under study. This could be due to a recent mutation which could have conferred an adaptive advantage of this mitotype over the other and could have erased past divergence during its dispersal. This M mitotype could also have been bottlenecked during the last glacial period leaving the entire territory for it to spread without any need of advantageous mutation (Śmietanka, Burzynski and Wenne, 2009). In any case, the *P. lacustris* situation demonstrates that two relatively divergent F mitochondrial genomes can function in concert with a single M genome.

4.6 Conclusion: evolution of M and F genomes in *Pyganodon*

Portions of mitochondrial genomes have long been used to estimate the evolutionary history of populations and species. All studies made to date using bivalve M mt genomes to assess population structure postulate that greater genetic diversity will lead to easier inference of fine-scale population differentiation (Śmietanka, Burzynski and Wenne, 2009). However, our study shows that this is not as straightforward as one might have previously thought and highlights that M mitochondrial genomes do not necessarily follow the same evolutionary paths as the conspecific F mitochondrial genomes. While the M and F genome's data revealed quite similar genetic patterns for *P. fragilis* and *P. cataracta*, the M genomes of *P. lacustris*, unlike their F counterpart, are almost homogeneous and *P. grandis* consists of an amalgam of divergent M and F lineages, which are not concordantly structured relative to geography.

Previous population genetic studies that used sequences from both M and F genomes have also yielded some paradoxical results. For example, in the

marine mussel *Musculista senhousia*, Passamonti (2007) found an inverted variability pattern with the F genome being more variable than the M genome. Śmietanka, Burzynski and Wenne (2009) also reported increased geographic differentiation in the F lineages, relative to the conspecific M lineages, of European *Mytilus* spp. By studying European mytilid mussel populations and by reinterpreting data from Riginos *et al.* (2004), Śmietanka, Burzynski and Wenne (2009) came to an appealing explanation for the disparity disparity among M and F genomes. They concluded that the overall pattern of polymorphism was consistent with periodic selective sweeps within the M lineage, leading to fixation of different variants in different populations, which was also consistent with more relaxed selection pressures in the M lineage in comparison with the F lineage. Our results in *Pyganodon* are particularly interesting in regard to this hypothesis because they highlight the possible/alternative scenarios that might occur under such a regime. Indeed, a selective sweep involving the M genome might have occurred recently in *P. lacustris* which has led to the loss of mitochondrial variation within populations and erased any geographical structure whereas the F genome data for this species reveal differentiation into two geographically structured clades. If we assume that selective sweeps happen periodically and repeatedly in M lineages and more rarely in F lineages, the spatial distribution of F haplotypes is more likely to represent ancient “history” whereas M haplotype distributions might reflect more recent population history (but without necessarily allowing inferences on fine-scale population structure if a sweep recently occurred). Following a sweep, the relaxation of selective constraints on the M molecule would rapidly re-establish a high level of polymorphism, which in return makes the sweep barely discernible.

4.7 Acknowledgements

We thank Dr Bob Krebs and an anonymous reviewer for suggestions that improved this article. I thank Philippe Boucher, AK Fritts, JA Bury and PL Phelps for their help collecting specimens. This work was supported by research grants from the National Sciences and Engineering Research Council (NSERC, to PU Blier) and the National Science Foundation (NSF DEB-0237175, to WR Hoeh), and the Minnesota Department of Natural Resources and U. S. Fish and Wildlife Service through Minnesota's State Wildlife Grants Program (SWG to BE Sietman). H. Doucet-Beaupré was financially supported by a NSERC scholarship and a NSERC fellowship, respectively. EG Chapman is supported by the Kentucky Agricultural Experiment Station State Project KY008043.

CHAPITRE V

POSITIVE SELECTION IN PATERNALLY INHERITED MITOCHONDRIAL GENOMES

Hélène Doucet-Beaupré¹, Philippe Doucet-Beaupré¹, Walter R. Hoeh², Pierre U. Blier¹,

¹Département de Biologie, Université du Québec à Rimouski, 300 Allée des Ursulines, Rimouski, Québec, Canada, G5L 3A1. (H. Doucet-Beaupré : conception du projet, analyse des données, rédaction de l'article; P. Doucet-Beaupré: programmation, traitement et analyse des données et conception des figures, aide à la rédaction; P. U. Blier : supervision et aide à la rédaction).

²Department of Biological Sciences, Kent State University, Kent, Ohio, 44242, USA. (W. R. Hoeh : aide à la rédaction).

Résumé

Dans certaines lignées de bivalves, il existe un système inhabituel d'héritabilité de l'ADN mitochondrial appelé la double héritabilité uniparentale (DUI). Dans ces organismes, des ADNmt distincts associés aux sexes parentaux coexistent, le génome maternel F et le génome paternel M. De nombreuses découvertes récentes permettent de mieux comprendre la structure et l'organisation du système de la double héritabilité uniparentale. Toutefois, les diverses forces évolutives qui ont pu générer et influencer la variation moléculaire des génomes M et F ne sont encore que peu étudiées. L'objectif spécifique de notre étude est de déterminer si les génomes mitochondriaux M ou F des espèces possédant la DUI sont soumis à une sélection positive dans le processus d'évolution, en utilisant un test branche-site pour une sélection positive. Nous avons utilisé 12 gènes concaténés codant pour des protéines de 29 génomes mitochondriaux de bivalve. La topologie des arbres phylogénétiques a été confirmée par quatre méthodes. Nous avons utilisé le programme CODEML de la suite PAML 4 pour effectuer les tests de sélection branche-site pour détecter des traces de sélection positive en utilisant des corrections multiples de Bonferroni. L'inférence Bayes empirique Bayes a été utilisée pour localiser les sites putatifs de sélection positive. Il existe de solides évidences de sélection positive sur l'ADN mitochondrial chez certaines lignées qui possèdent la DUI. Chez les Unionoidae, nous avons détecté de la sélection positive significative sur 112 sites à travers tous les gènes du génome M et sur 30 sites sur les génomes F. Chez les Veneridae, le test détecte une importante sélection positive sur 38 sites dans le génome F et seulement 12 sites dans le génome M. Les sites sélectionnés positivement sont essentiellement localisés dans deux gènes, *nd4* et *nd4l*, et sont probablement associés à la réorganisation génique observée dans le génome F. Chez les Mytilidae, le test n'a pas détecté de sélection positive. Selon la présente étude, la sélection naturelle affecterait différemment chaque lignée de bivalves qui possède la DUI.

5.1 Abstract

In some bivalve lineages, an unusual mitochondrial DNA inheritance system termed doubly uniparental inheritance occurs. In those organisms, distinct gender-associated mtDNAs coexist; the maternally inherited F genomes and the paternally inherited M genomes. Despite many recent insights into the structure and organization of the DUI system, much remain to be learned about the evolutionary forces that could have generated and influenced the molecular variation of the M and F genome. The specific goal of our study is to identify whether the M or the F mt genomes of DUI species is subject to positive selection in the process of evolution by using a branch-site test for positive selection. We used 12 concatenated protein-coding genes from 29 bivalve mitochondrial genomes. Phylogenetic tree topologies were confirmed by four methods. Multiple Bonferroni correction improved branch-site test for positive selection using CODEML program from the PAML 4 suite. Bayes Empirical Bayes inference was used to localize putative positively selected sites. There is strong evidence of positive selection on mitochondrial DNA of certain lineages that possesses DUI. In Unionoidae, we detected significant positive selection on 112 sites across all genes of the M genome and on 30 sites on F genomes. In Veneridae, the test detected significant positive selection on 38 sites in the F genome and only 12 sites in the M genome. Putative positively selected sites, are essentially localized in two genes ND4 and ND4L and is likely associated to genome reorganization in the F genome instead of adaptive evolution associated to DUI. In Mytilidae, the test didn't detect positive selection. The evolutionary fate of DUI appears different according to each bivalve lineage.

5.2 Introduction

Almost all animal cells contain inheritable genetic material, independent of the nuclear chromosomes: the mitochondrial DNA (mtDNA) genomes. Located in the mitochondrial matrix, the metazoan mtDNA span usually about 14-16 kilobases (kb) and encodes thirteen peptides pivotal for cellular energy metabolism [Boore, 1999; Burger, Gray and Lang, 2003]; but see Birky, 2001 for exceptions]. An elaborate and complex cellular machinery and quantity of nuclear encoded factors ensure expression, replication and maintenance of mtDNA (Shadel, 2008). Inheritance of mtDNA is essentially non-mendelian; recombination through meiosis is lacking and uniparental inheritance from maternal is the general rule in animal species (Birky, 1995; Birky, 2001). In contrast to nuclear genome, in which there are only two copies of each gene per cell, mtDNA is maintained at a high copy number [between 10^3 – 10^4 copies per somatic cell, organized as 2-10 copies per organelle Nass, 1969; Satoh and Kuroiwa, 1991]. These peculiar and still incompletely understood features coupled with the successive cell divisions that represent sequential bottlenecks for the mitochondrial population (Birky, 2001; Rand, Clark and Kann, 2001; Shoubridge and Wai, 2008) contributed to preserve a homoplasmic state of the mtDNA, i.e. all copies of mtDNA in each cell typically have at birth identical DNA sequences.

However, high rates of nucleotide substitution (Moritz, Dowling and Brown, 1987; Pesole *et al.*, 1999) and limited DNA repair mechanisms leave mt genes susceptible to rapid evolution through random drift or natural selection (Meiklejohn, Montooth and Rand, 2007; Dowling, Friberg and Lindell, 2008). As a result, mtDNA is highly polymorphic and have been used widely to study population genetic structure and phylogenetic relationships among animals. MtDNA is highly polymorphic at the population level, but this

is not the case at the individual level. The occurrence of more than one mtDNA sequence variant in a single individual (heteroplasmy) is relatively rare (Jenuth *et al.*, 1996; Stewart *et al.*, 2008b) and generally resolved to homoplasmy within a few generations (Bendall, Macaulay and Sykes, 1997; Chinnery, 2002).

The most noteworthy exception to this axiomatic rule is found in bivalve lineages (i.e., the orders Mytiloida, Unionoida and Veneroida), in which two highly divergent mitochondrial variants are inherited: one through females (F) and the other through males (M). This unusual system is known as doubly uniparental inheritance (DUI) as both sexes inherit uniparentally their mitochondria [for reviews see Breton *et al.*, 2007 and Passamonti and Ghiselli, 2009]. M and F mitochondrial genome can exhibit up to 52% sequence divergence and coexist inside an individual (Doucet-Beaupré *et al.*, 2010).

Notwithstanding many recent insights into the structure and organization of the DUI system, much remains to be learned about many aspects of M and F genome evolution and the evolutionary implications of DUI. Since the discovery of the system, it is unclear whether a selective advantage favored the retention of DUI in bivalves but the idea that DUI may have allowed adaptive evolution in paternally inherited genomes have always been omnipresent (Skibinski, Gallagher and Beynon, 1994; Gemmell and Braisher, 2001; Selosse, Albert and Godelle, 2001; Zeh and Zeh, 2005). Intra- and interspecific comparisons suggest that the M genome is evolving more rapidly than the F genome (Skibinski, Gallagher and Beynon, 1994; Rawson and Hilbish, 1995; Stewart *et al.*, 1995; Liu, Mitton and Wu, 1996; Hoeh, Stewart and Guttman, 2002; Krebs, 2004) but the reasons for this accelerated

evolution remain obscure. Recent population genetics studies based on M and F inherited genes found patterns best explained by relaxation of selective constraints and episodes of adaptive evolution occurring more frequently within the M lineage than the F lineage (Smietanka, Burzynski and Wenne, 2009). Nevertheless, there is still no conclusive evidence of adaptive evolution in the paternally transmitted genome.

Positive selection (often also referred to as adaptive molecular evolution) is difficult to detect because it often operates episodically on a few amino acid sites. There are two classes of statistical methods designed to detect traces of positive selection on molecular data. The first class of method is the tests of neutrality [for example, McDonald and Kreitman, 1991]. Used since a long time by population geneticists, they are adapted for within-species polymorphism. Those tests concern variations that are not yet fixed in time. Statistically, they compared the observed distribution data to a neutral predicted data set based on the neutral theory of evolution. They are unsuitable for multilocus data and for interspecific data. The alternative to tests of neutrality is the comparison of substitution rate in coding sequence. This method uses the estimation of D_n (non synonymous substitution rate) and D_s (synonymous substitution rate; named synonymous because they don't modify the encoded amino acid sequences and are considered *a priori* as neutral). The ratio D_n/D_s (or ω) has become a standard measure of selective pressure $\omega \approx 1$ signifies neutral evolution, $\omega < 1$: negative selection and $\omega > 1$: positive selection.

Early statistical tests based on a traditional pairwise approach, simply counted the rate of non-synonymous (dN) and synonymous (dS) substitution

between two sequences averaged over-all codons but were heavily biased due to simplistic assumptions (Yang, 2002). With this approach, the signal is usually masked by negative selection and detecting positive selection is difficult (Zhang, Nielsen and Yang, 2005). Most of the previous analyses of M inherited genes depended upon statistical tests based on a traditional pairwise approach (Stewart *et al.*, 1995; Hoeh *et al.*, 1996a; Stewart *et al.*, 1996; Quesada, Wenne and Skibinski, 1999). Alternatively, Markov processes can be used for the statistical modeling of evolution. Models, more or less complex, are used to describe nucleotide, amino acid or codon changes in evolutionary time (Anisimova and Liberles, 2007). To apply those models to a data set a probabilistic inference method has to be used. It could be either maximum likelihood or Bayesian inference. An aligned data set, a phylogenetic tree and a model on a matrix are used to estimate the probability to obtain the data set according to the specified tree and model. With this approach, it's then possible to compare the probability of a model which includes positive selection to the probability of a model which not allows for positive selection. By comparing the probabilities of the two models, the best-fit model of evolution for a particular data set can be discriminate. Several models were developed; some are simples, including only few parameters and considering evolutionary processes as homogenous [for example Nei and Gojobori, 1986 and Yang, 1993]. More sophisticated models have been developed and now incorporate the heterogeneity in the evolutionary process along sequences (Yang *et al.*, 2000; Whelan and Goldman, 2001). The models gained in complexity and in parameters. For example, it's possible to let ω varies between the different sites of the molecule (site-models) or it's possible to let branches of the tree evolved independently (branch-models) and there are some models that can do both (branch-site models).

Currently, many different methods exist to detect natural selection acting at certain amino acid sites [Anisimova and Liberles, 2007 for review]. The specific goal of our study is to identify whether the M or the F mt genomes of DUI species is subject to positive selection in the process of evolution by using an ameliorated branch-site test of positive selection.

5.3 Materials and methods

5.3.1 Sequence alignment

Amino acids from 29 complete bivalve mitochondrial genomes and those from five outgroup species (gastropod *Aplysia californica* [NC_005827], cephalopod *Octopus vulgaris* [NC_006353], scaphopod *Graptacme eborea* [NC_006162], polyplacophoran *Katharina tunicata* [NC_001636], polychaete *Platynereis dumerii* [NC_000931]) were aligned (Table 5.1). We used both Clustal W (Larkin *et al.*, 2007) and Dialign version 2.2.1 (Subramanian, Kaufmann and Morgenstern, 2008) for the alignments, with subsequent manual adjustments, and the amino acid alignment was used as a template to align the corresponding codons. Regions of ambiguous alignment were excluded prior to the phylogenetic analyses. The analyzed matrices had either 7,704 nucleotide positions or 2,568 amino acid positions, and these files are available from the authors.

Table 5.1 Names and GenBank accession numbers of the 34 species used in our evolutionary analysis of bivalve mitochondrial genomes.

Species	Gender	GenBank accession Numbers
Bivalvia		
Autolamellibranchiata		
Unionoida		
<i>Venustaconcha ellipsiformis</i>	F	FJ809753
<i>Venustaconcha ellipsiformis</i>	M	FJ809752
<i>Pyganodon grandis</i>	F	FJ809754
<i>Pyganodon grandis</i>	M	FJ809755
<i>Inversidens japonensis</i>	F	AB055625
<i>Inversidens japonensis</i>	M	AB055624
<i>Quadrula quadrula</i>	F	FJ809750
<i>Quadrula quadrula</i>	M	FJ809751
<i>Lampsilis ornata</i>	F	NC_005335
<i>Hyriopsis cumingii</i>	F	NC_011763
Veneroida		
<i>Venerupis philippinarum</i>	F	NC_003354
<i>Venerupis philippinarum</i>	M	AB065374
<i>Acanthocardia tuberculata</i>	—	NC_008452
<i>Sinonovacula constricta</i>	—	NC_011075
<i>Hiatella arctica</i>	—	NC_008451
Pteriomorpha		
Mytiloida		
<i>Mytilus trossulus</i>	F	DQ198231
<i>Mytilus trossulus</i>	M	DQ198225
<i>Mytilus edulis</i>	F	NC_006161
<i>Mytilus edulis</i>	M	AY823624
<i>Mytilus galloprovincialis</i>	F	NC_006886
<i>Mytilus galloprovincialis</i>	M	AY363687
Ostreoida		
<i>Crassostrea hongkongensis</i>	—	NC_011518
<i>Crassostrea gigas</i>	—	NC_001276
<i>Crassostrea virginica</i>	—	NC_007175
Pectinoida	—	NC_009687
<i>Argopecten irradians</i>		
<i>Placopecten magellanicus</i>	—	NC_007234
<i>Mizuhopecten yessoensis</i>	—	NC_009081
<i>Chlamys farreri</i>	—	NC_012138
<i>Mimachlamys nobilis</i>	—	NC_011608
Out-group		
<i>Aplysia californica</i> (Gastropoda)	—	NC_005827
<i>Graptacme eborea</i> (Scaphopoda)	—	NC_006162
<i>Katharina tunicata</i> (Polyplacophora)	—	NC_001636
<i>Octopus vulgaris</i> (Cephalopoda)	—	NC_006353
<i>Platynereis dumerii</i> (Polychaeta)	—	NC_000931

5.3.2 Phylogeny reconstruction

Phylogenetic trees for the Bivalvia, using Bayesian inference (BI), maximum likelihood (ML) and maximum parsimony (MP), were constructed using concatenated nucleotide and amino acid sequences from 12 protein-coding genes (we excluded *atp8* due to alignment issues and its apparent absence in multiple bivalve species). The methodological details for the phylogenetic reconstruction can be found in Doucet-Beaupré *et al.* (2010).

The codon- and amino acid-based BI analyses were conducted with Mr. Bayes version 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The codon-based analysis invoked the M3 model (Yang *et al.*, 2000) with two simultaneous runs of 5 million generations each (a total of 50,000 saved trees/run). The amino acid-based BI analysis invoked the variable rate "glorified GTR model" [see the MrBayes manual; Ronquist, Huelsenbeck and van der Mark, (2005)] with two simultaneous runs of 2.9 million generations each (a total of 29,000 saved trees/run). Both sets of BI analyses reached convergence (average standard deviation of the split frequencies was <0.01) and the burnin for each set was determined by reference to the log probability of observing the data x generation plot (codon-based BI run burnin = 4 million generations [= the last 10,000 trees/run saved contributed to the majority-rule tree], amino acid-based BI run burnin = 1.9 million generations [= the last 10,000 trees/run saved contributed to the majority-rule tree]).

Codon and amino acid-based ML analyses were conducted with Garli version 0.96 (Zwickl, 2006). The M3 model was used in the codon-based ML analysis, which was set to use the observed nucleotide frequencies at each

codon position separately. A non-parametric bootstrap (Felsenstein, 1985) analysis was performed, using 300 replicates, to assess nodal support for the codon analysis-based trees. The program ProtTest (<http://darwin.uvigo.es/software/prottest.html>) was used to evaluate the best amino acid model for our data from those models available in Garli. Both the Akaike information criterion and Bayesian information criterion selected the WAG+F model (Whelan and Goldman, 2001) as the best available model which was therefore used in the ML amino acid analysis. A non-parametric bootstrap was performed, using 600 replicates, to assess nodal support for the amino acid analysis-based trees.

Maximum parsimony analyses were conducted with PAUP* (Swofford, 2002). The nucleotide-based MP analysis utilized equally weighted transversion parsimony (= only purines vs. pyrimidines were coded) and 1000 random addition runs for estimating the most parsimonious tree. A non-parametric bootstrap transversion parsimony analysis was run (with 1000 replicates) using 10 random addition runs per replicate. The amino acid-based MP analysis was carried out with equal weighting and 1000 random addition runs were used to estimate the most parsimonious tree. Lastly, an equally weighted parsimony, non-parametric, bootstrap analysis was run on the amino acid matrix (with 1000 replicates) using 10 random addition runs per replicate.

5.3.3 Analysis of selective constraints

To analyze selective constraints acting on the bivalve mitochondrial genome along certain lineages, and to infer possible amino acid sites under positive selection, we use the improved branch-site test for positive selection

(BST) (Wang and Guo, 2004; Zhang, Nielsen and Yang, 2005; Anisimova and Liberles, 2007); followed by Bayes Empirical Bayes (BEB) inference (Yang, Wong and Nielsen, 2005). Both methods are implemented in CODEML of the PAML 4 suite (Yang, 2007). To set up the likelihood ratio test (LRT), which include the BST using a branch-site codon model, we first estimated branch lengths along the previously determined tree in an initial pass using the neutral (M1a) codon substitution models. This allowed avoiding the optimization of 65 length estimates simultaneously with branch-site model parameters. Indeed, we noticed that the convergence of a completely length-unconstrained branch-site model poses many difficulties because of the strong cross-correlation in the foreground branch between arbitrarily small lengths and arbitrarily large ω . We, therefore, chose to trade numerical difficulties for a loss of power in the BST by fixing branch lengths to their prior point estimate and thereby degrading their correlation with the foreground and background ω parameters. Each optimization is repeated using two different schemes of branch length perturbations (simultaneous or one at a time) and in each case 4 runs with different initial random lengths and model parameter values are performed from which we kept consensual runs with the lowest log-likelihood. With these estimates established, we then proceeded with several LRTs where the alternative and null hypotheses for each LRT are set up following Yang (1998); Anisimova, Bielawski and Yang, (2001) between two M1a+M2a branch-side models. The foreground branch is in turn *a priori* chosen to be the branch spawning the male or female Unionoidae, Mytilidae or Veneridae clades for a total of 6 LRTs. In the null hypotheses, background branches follow the M1a model whereby only negative ($\omega_0 < 1$) and neutral ($\omega_1 = 1$) selection is allowed while the foreground branch instead follows the M2a model whereby an additional ω -class is included ($\omega_2 \geq 1$) which we constrain at $\omega_2 = 1$. Since ω_0 on the

foreground branch is the same as the ω_0 on background branches, the maximum likelihood estimate will seek to best explain elevated rates of non-synonymous substitutions along the foreground branch as a relaxation of selective constraints by elevating the proportion p_1 of neutrally selected site instead of positive selection. In the alternative hypotheses, $\omega_2 \geq 1$ is left unconstrained and therefore estimated. Each codon site has a probability p_0 , p_1 and p_2 of belonging to ω_0 , ω_1 and ω_2 rate ratio classes, respectively, with p_2 fixed by $p_0 + p_1 + p_2 = 1$ and the rate ratio of a given site is therefore the weighted average over ω -classes. The statistic $-2\Delta l$ is then compared against the χ_1^2 distribution with 1 degree of freedom¹ to determine the significance of the LRT. Although the null hypothesis has one of its parameters ($\omega_2 = 1$) fixed at a boundary of the parameter space of the alternative hypothesis ($\omega_2 \geq 1$), comparing against χ_1^2 , instead of against a 50:50 mixture between a point mass at 0 and χ_1^2 , as one would conventionally choose, makes the test more conservative against regularity and assumptions violation. We test 6 hypotheses over the same dataset and therefore reject each individual null hypothesis at the Bonferonni corrected significance level $\alpha_{BC} = 0.05/6 = 0.008$. Finally, we proceed with BEB inference wherever the LRT suggests positive selection has occurred along any of the six previous lineages. To infer the probability of belonging to the $\omega_2 \geq 1$ class of positively selected sites, we chose the truncated uniform prior $\omega_2 \sim U(1,31)$ discretized into 50 equally spaced bins. All the above codon models used the FMutSel-F

¹ The alternative hypothesis parameters are p_0 , p_1 , ω_0 , ω_2 , κ and 3 stationary nucleotides frequencies π_i from FMutSel-F (the fourth one is fixed by normalization) for a total of 8 degrees of freedom. The null hypothesis fixes $\omega_2 = 1$ for a total of 7 degrees of freedom. One would therefore conventionally expect the LRT to asymptotically behave as $-2\Delta l \sim \frac{1}{2} (\chi_1^2 + \delta_0)$ with δ_0 the point-mass distribution at 0.

substitution model (Yang and Nielsen, 2008) to accommodate the branch-site model against the bias introduced into stationary codon and nucleotide frequencies by the combined effect of mutation and selection (Albu *et al.*, 2008). Since FMutSel-F determines a unique codon fitness for each 62 codons of invertebrate mtDNA across all 34 species based on observed frequencies, the fitness therefore assumes a unique effective population across the whole tree. Finally each branch-site model optimization, i.e. 2 optimization per LRT, is repeated 8 times with different initial ω values and discrete difference parameters from which we kept the lowest consensual log-likelihood between two or more runs.

5.4 Results

We examined the possibility that positive selection acts on maternally and/or paternally inherited mitochondrial genes of bivalve species possessing DUI (Figure 5.1). According to our analyses, there is strong evidence of positive selection in mitochondrial DNA of some of our *a priori* targeted lineages (Table 5.2, Figures 5.2–5.4).

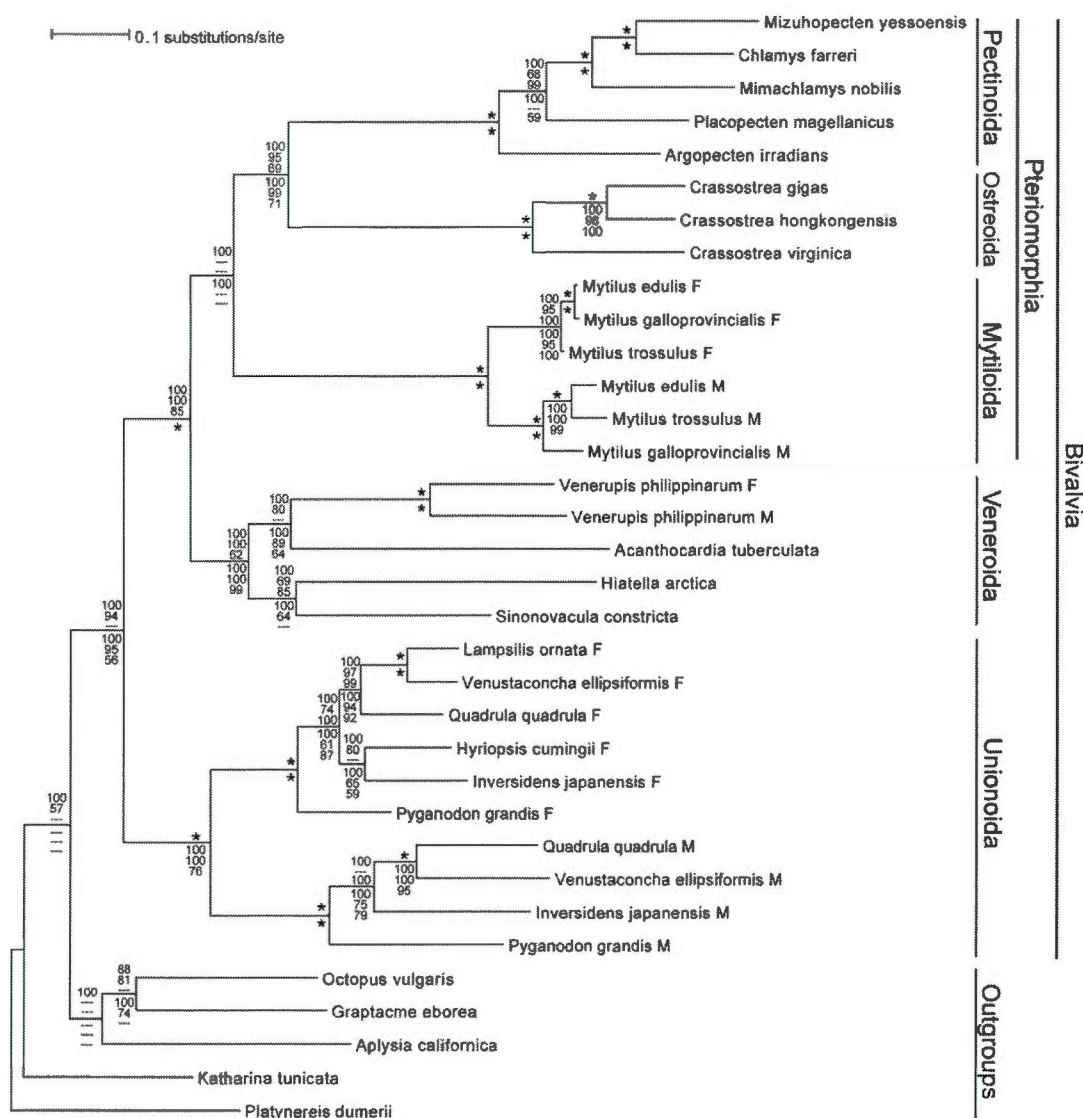


Figure 5.1 DUI phylogeny. The mitochondrial genome evolutionary relationships for the DUI-containing taxa showing that the male (M) and female (F) genomes in the freshwater mussels, the Unionidae, have been stable whereas there is evidence for occasional masculinization events (i.e. F genomes giving rise to M genomes) in the families Veneridae and Mytilidae. Topology and branch lengths are based on analysis of 199 amino acids from *cox1* and posterior probabilities ($\times 100$) > 90 are displayed. The taxonomic and sex-specific transmission affiliations of the individual sequences are indicated at the right. The Mr. Bayes analysis (Huelsenbeck and Ronquist, 2001) used the Jones, Taylor and Thornton (JTT) model

of amino acid substitution (Jones, Taylor and Thornton, 1992) and was run for 10 million generations with eight chains. Overall, 10 000 trees were saved during the course of the analysis but only the last 9000 were used to construct the consensus tree (i.e. the first 1000 trees were discarded as burn-in). The original nucleotide sequences (species and accessions numbers) were translated with the *Drosophila* mitochondrial genetic code and used according to Refs (Beagley, Taylor and Wolstenholme, 1997; Hoeh *et al.*, 1998; Hoeh, Stewart and Guttman, 2002; Boore, Medina and Rosenberg, 2004; Mock *et al.*, 2004; Riginos *et al.*, 2004; Mizi *et al.*, 2005; Breton *et al.*, 2006; Doucet-Beaupré *et al.*, 2010).

Table 5.2. Summary of branch-site test for positive selection along the 6 *a priori* selected lineages.

Scheme	np	Point estimates	log likelihood /
Male Unionoidae foreground			
M2a null hypothesis	3	$p_0=0.81996, p_1=0.06506,$ $(p_2=0.11498), \omega_0=0.05718,$ $(\omega_1=\omega_2=1)$	-178562
M2a alternative hypothesis	4	$p_0=0.76146, p_1=0.06551,$ $(p_2=0.17303), \omega_0=0.05854, (\omega_1=1),$ $\omega_2=21.31042$	-178490
			<hr/> -2 Δl = 144 (p-value << 10 ⁻⁵)
Female Unionoidae foreground			
M2a null hypothesis	3	$p_0=0.86177, p_1=0.06697,$ $(p_2=0.07126), \omega_0=0.05733,$ $(\omega_1=\omega_2=1)$	-178601
M2a alternative hypothesis	4	$p_0=0.84014, p_1=0.06683,$ $(p_2=0.09303), \omega_0=0.05805, (\omega_1=1),$ $\omega_2=14.73065$	-178568
			<hr/> -2 Δl = 66 (p-value << 10 ⁻⁵)
Male Veneridae foreground			
M2a null hypothesis	3	$p_0=0.87296, p_1=0.06914,$ $(p_2=0.05790), \omega_0=0.05747,$ $(\omega_1=\omega_2=1)$	-178610
M2a alternative hypothesis	4	$p_0=0.87736, p_1=0.06963,$ $(p_2=0.05301), \omega_0=0.05776, (\omega_1=1),$ $\omega_2=8.62523$	-178593
			<hr/> -2 Δl = 34 (p-value << 10 ⁻⁵)

Scheme	np	Point estimates	log likelihood /
Female Veneridae foreground			
M2a null hypothesis	3	$p_0=0.85305, p_1=0.06734,$ $(p_2=0.07961), \omega_0 = 0.05727,$ $(\omega_1=\omega_2=1)$	-178575
M2a alternative hypothesis	4	$p_0=0.85409, p_1=0.06784,$ $(p_2=0.07807), \omega_0=0.05767, (\omega_1=1),$ $\omega_2=8.32673$	-178537
			<hr/> -2 Δl = 76 (p-value << 10 ⁻⁵)
Male Mytilidae foreground			
M2a null hypothesis	3	$p_0=0.92361, p_1=0.07432,$ $(p_2=0.00207), \omega_0=0.05792,$ $(\omega_1=\omega_2=1)$	-178636.1
M2a alternative hypothesis	4	$p_0=0.92296, p_1=0.07392,$ $(p_2=0.00312), \omega_0=0.05787, (\omega_1=1),$ $\omega_2=5.53543$	-178634.6
			<hr/> -2 Δl = 3.0 (p-value = 0.083)
Female Mytilidae foreground			
M2a null hypothesis	3	$p_0=0.92019, p_1=0.07392,$ $(p_2=0.00589), \omega_0=0.05789,$ $(\omega_1=\omega_2=1)$	-178635.5
M2a alternative hypothesis	4	$p_0=0.92175, p_1=0.07376,$ $(p_2=0.00449), \omega_0=0.05783, (\omega_1=1),$ $\omega_2=6.08919$	-178632.3
			<hr/> -2 Δl = 6.4 (p-value = 0.011)

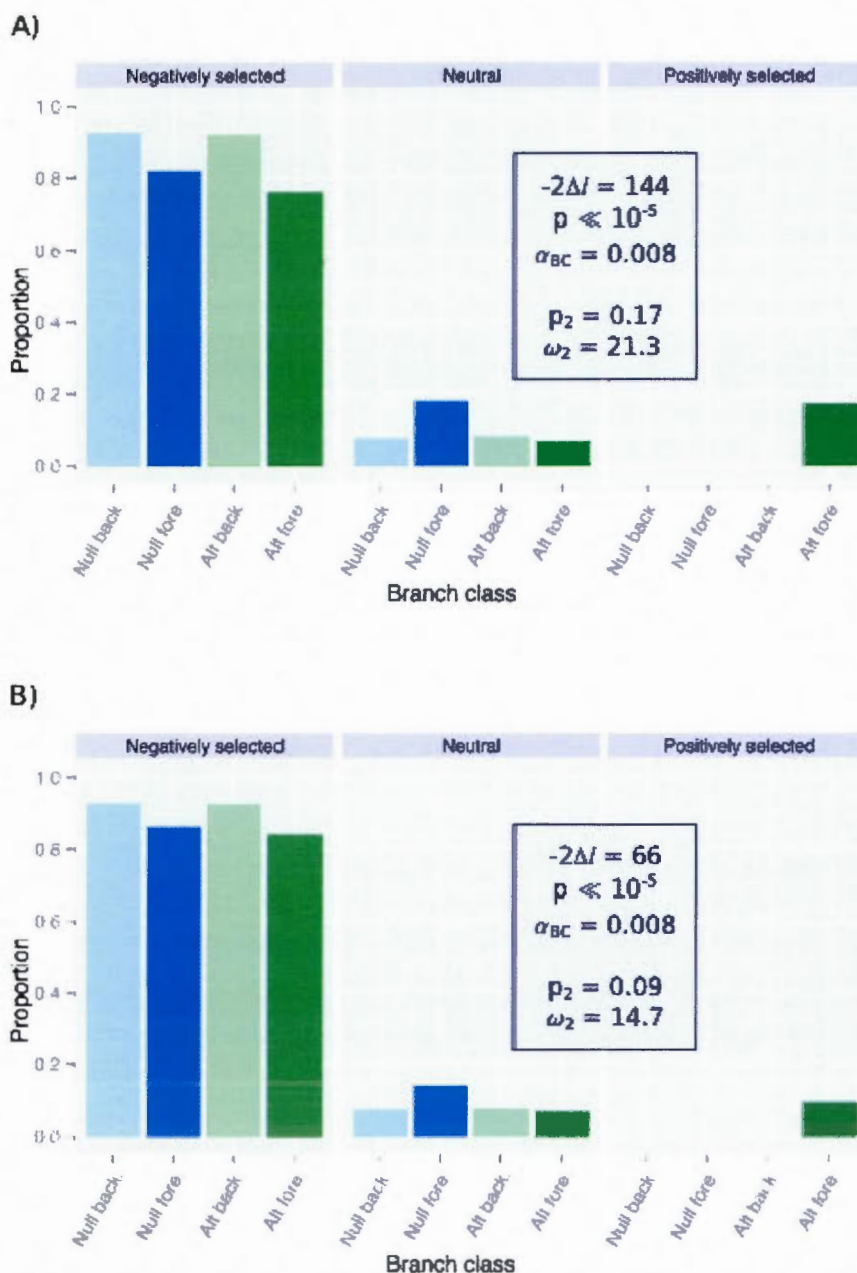


Figure 5.2. Proportion of sites in ω classes (negatively selected, neutral or positively selected) according to their branch classes (background (back) vs. foreground (fore)) and null vs. alternative (alt) hypothesis in the unionoid genome (A: Male transmitted genome; B: Female transmitted genome). $-2\Delta l$: likelihood ratio test; p : χ^2 probability; α_{BC} : Bonferroni corrected; p_2 : estimated proportion of sites in class $\omega > 1$. ω_2 : maximum likelihood point estimate for ω_2 on the foreground branch for the fraction of sites under positive selection.

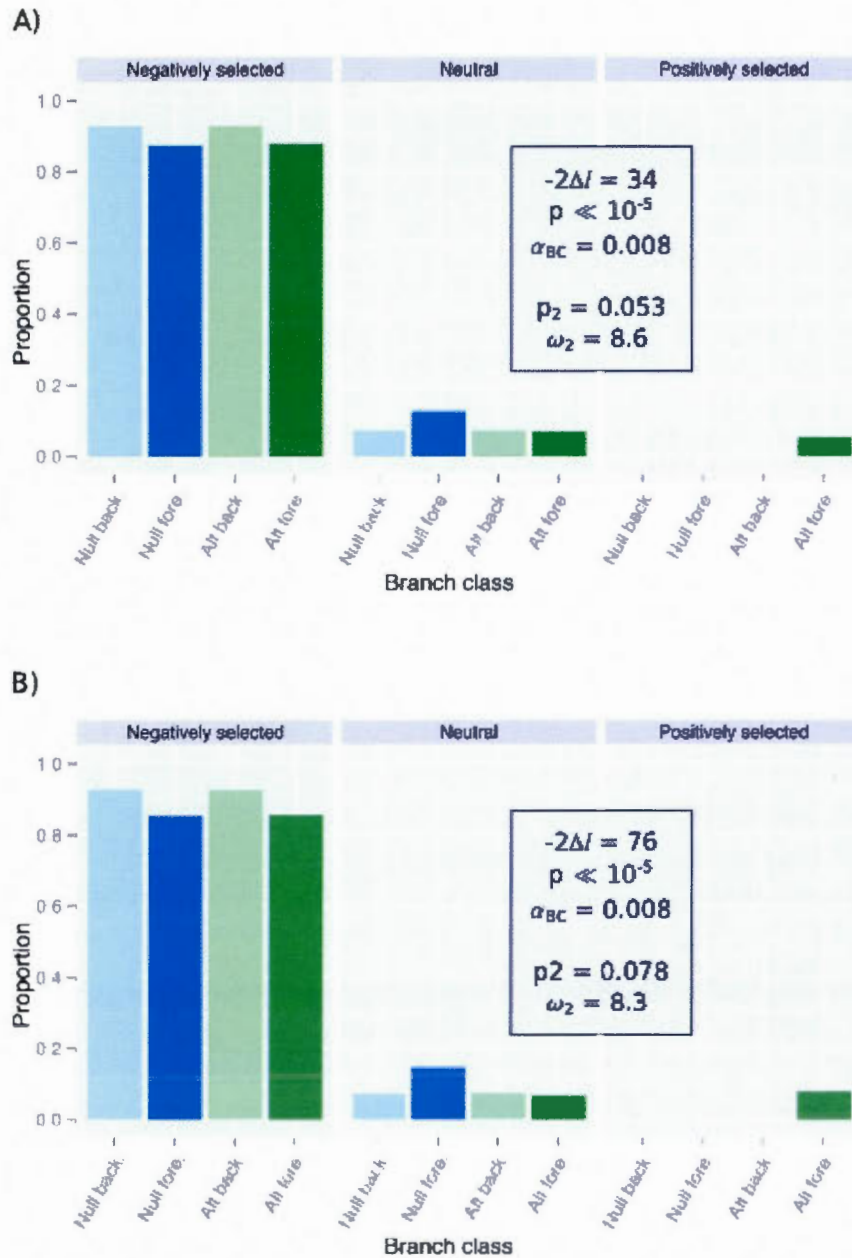


Figure 5.3. Proportion of sites in ω classes (negatively selected, neutral or positively selected) according to their branch classes (background (back) vs. foreground (fore)) and null vs. alternative (alt) hypothesis in the veneroid genome (A: Male transmitted genome; B: Female transmitted genome). $-2\Delta l$: likelihood ratio test; p : χ^2 probability; α_{BC} : Bonferroni corrected; p_2 : estimated proportion of sites in class $\omega > 1$. ω_2 : maximum likelihood point estimate for ω_2 on the foreground branch for the fraction of sites under positive selection.

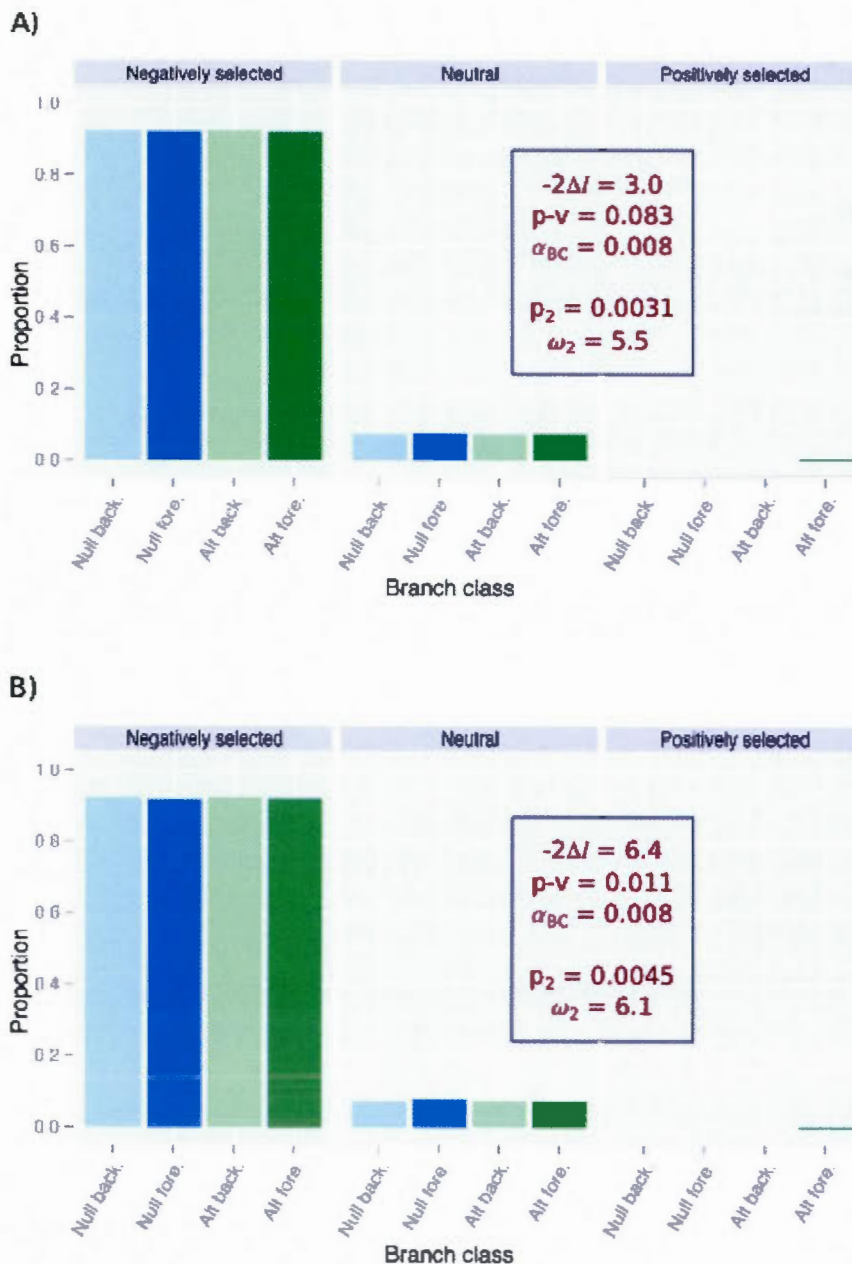


Figure 5.4. Proportion of sites in ω classes (negatively selected, neutral or positively selected) according to their branch classes (background (back) vs. foreground (fore)) and null vs. alternative (alt) hypothesis in the mytiloid genome (A: Male transmitted genome; B: Female transmitted genome). $-2\Delta l$: likelihood ratio test; p : χ^2 probability; α_{BC} : Bonferroni corrected; p_2 : estimated proportion of sites in class $\omega > 1$. ω_2 : maximum likelihood point estimate for ω_2 on the foreground branch for the fraction of sites under positive selection.

When testing the hypothesis that codon sites within paternally inherited genomes of Unionoidae have experienced adaptive evolution rather than selection constraints relaxation, significant results were obtained for several codons within all genes examined (Figure 5.2A). The BST for paternally inherited mitochondrial genome of Unionoidae yielded a LRT statistic $-2\Delta/ = 144$ (p-value $\ll 10^{-5}$) and a proportion of sites $p_2 = 17.3\%$ under the rate ratio $\omega_2 = 21.3$. In contrast with this point estimate for ω_2 , the marginalized BEB posteriori (Figure 5.5) gives an improved summary and visual representation of the uncertainty over ω_2 whereby we find a BEB inferred positive selection rate ratio $\omega_2 = 4.9 \pm 0.67$ (average $\pm \sigma$), much lower than the point estimate. Of the 2568 codon sites analyzed (Figure 5.6 and 5.7), 112 of them were identified as sites of positive selection with strong BEB posterior probability $\Pr(\omega_2) > 95\%$ and 303 with $\Pr(\omega_2) > 75\%$ (see appendice C.1 for Supplementary Figure). Similarly, we also detected significant positive selection on sites of the F genome of Unionoidae (Figure 5.2B) with LRT statistic of $-2\Delta/ = 66$ (p-value $\ll 10^{-5}$), but in point estimated proportion $p_2 = 9.3\%$ and rate ratio $\omega_2 = 14.7$. The BEB inferred positive selection rate ratio rather gives $\omega_2 = 4.7 \pm 3.4$ (Figure 5.5). Here we see that while the point estimate for ω_2 in the F genome is much lower than what occurs in M genome, the BEB estimate for the F genome is close to the one of the M genome but with larger uncertainty. Among those sites in the F genome (Figure 5.6 and 5.7), BEB analysis identified 30 amino acid sites with strong support $\Pr(\omega_2) > 95\%$ and 84 amino acid sites with support $\Pr(\omega_2) > 75\%$ (see appendice C.2 for Supplementary Figure).

When looking at the ω -class probability distribution over sites across all 12 genes examined (Figures 5.6 and 5.7), in both the M genome and the F genomes of Unionoidae, sites with high posterior probability of evolving under

positive selection are generally distributed across all 12 genes with no clear pattern.

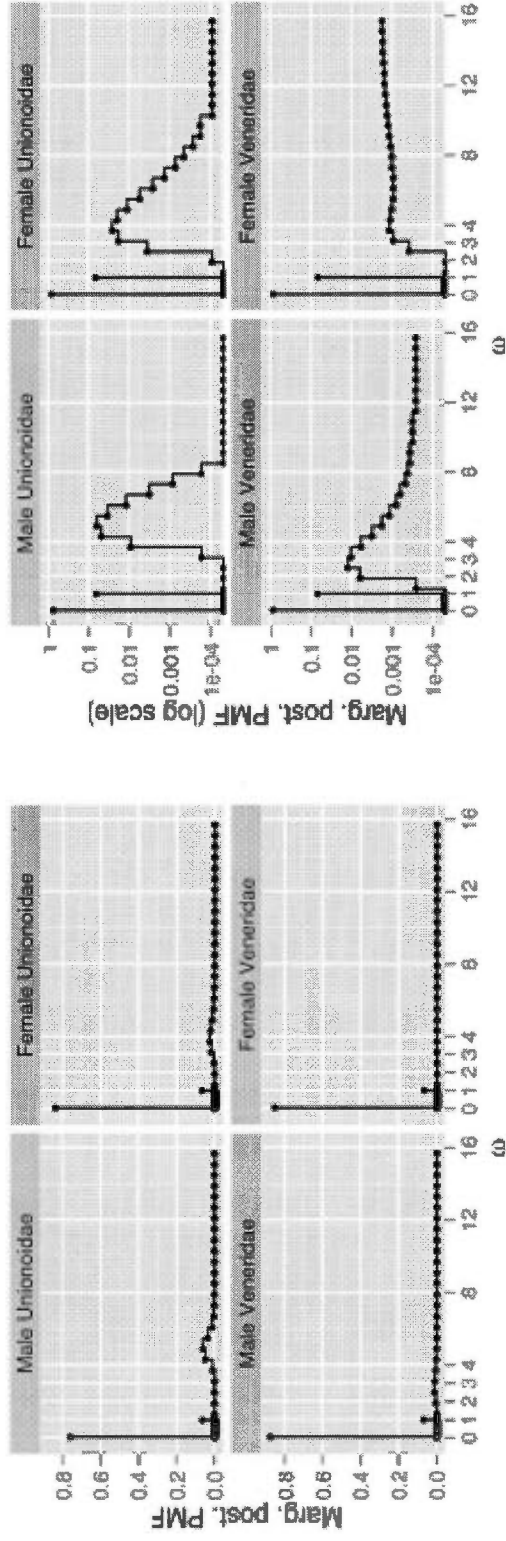


Figure 5.5. Posterior probability mass functions (PMF) of ω marginalized over foreground branch sites and whole background for targeted lineages where likelihood ratio test gives significant results of positive selection. Dots show the discretization $\omega_0 \sim U(0,1)$, $\omega_1 \sim \delta(1)$, $\omega_2 \sim U(1,31)$ and steps are guide to the eye. The rightmost figure show the same PMFs on a logarithmic scale to emphasize the small but non-zero probability mass of sites to be under positive selection.

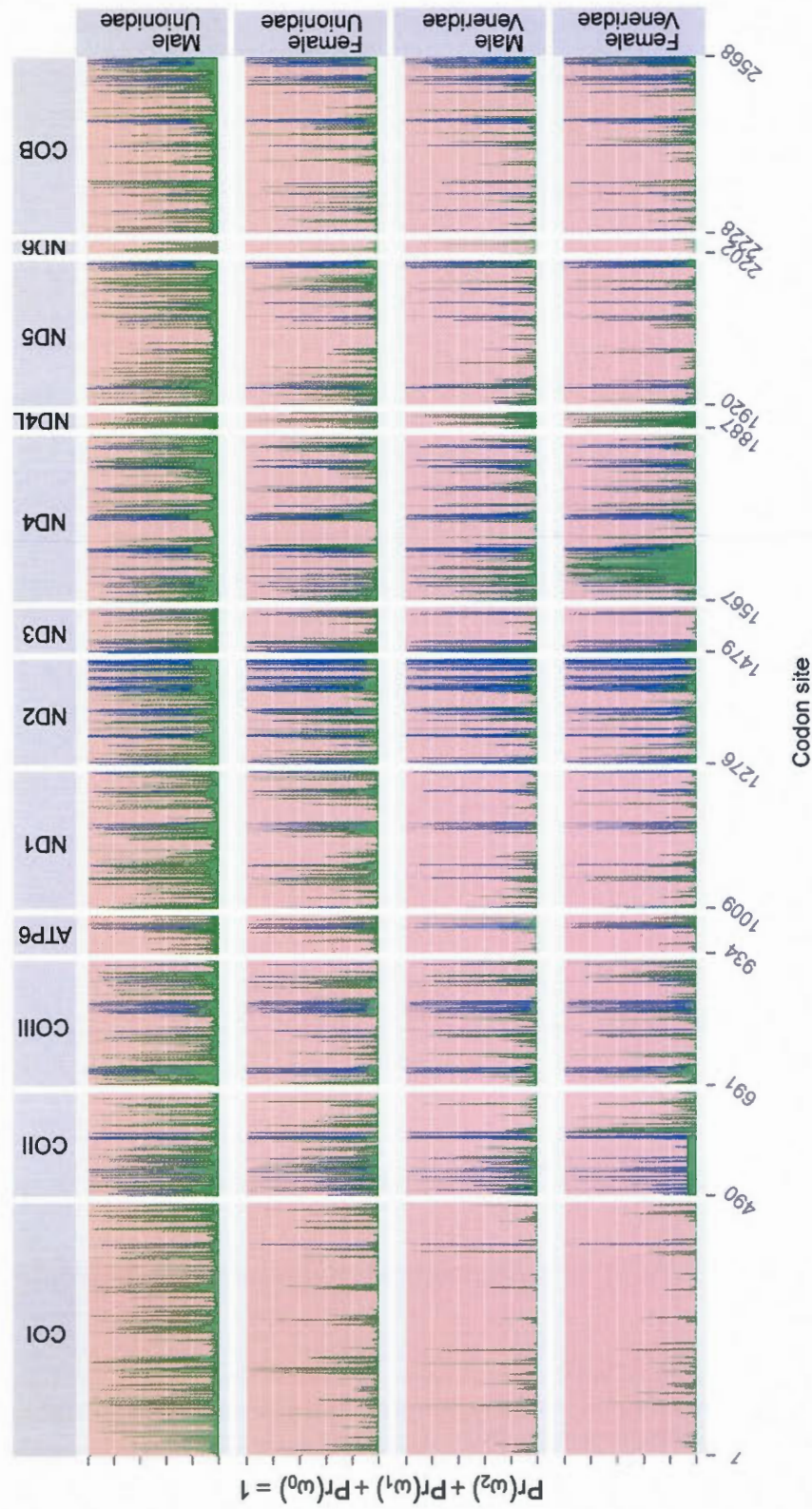


Figure 5.6. Bayesian empirical Bayes (BEB) probability for each site to be either negatively selected ($\Pr(\omega_0)$; red), neutral ($\Pr(\omega_1)$; blue) or positively selected ($\Pr(\omega_2)$; green) in targeted lineages where likelihood ratio test give significant results of positive selection.

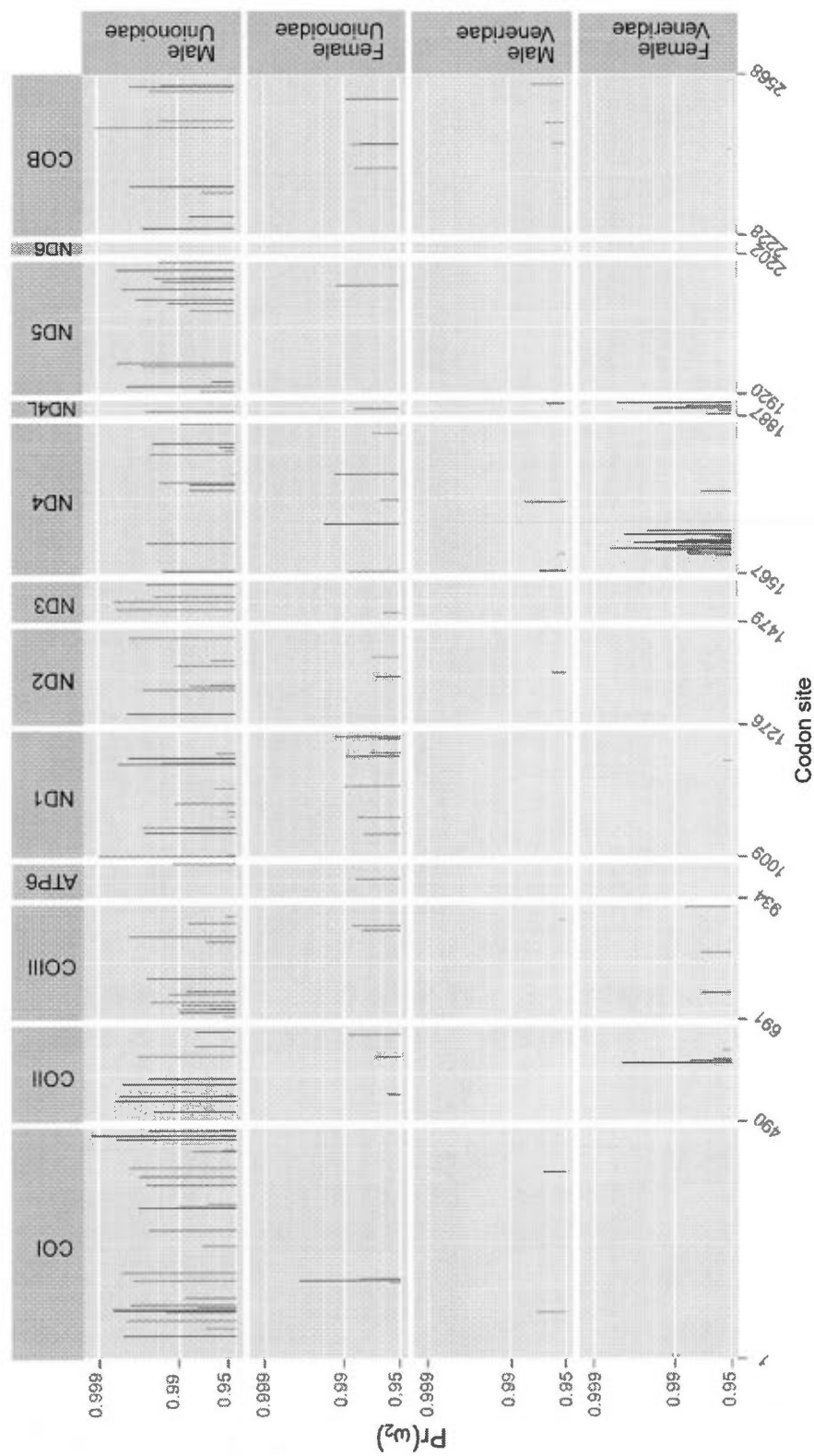


Figure 5.7. Bayesian empirical Bayes (BEB) probability for each site to be positively selected ($Pr(\omega_2) \geq 0.95$) in targeted lineages where likelihood ratio test give significant results of positive selection.

In Veneridae, evidence of positive selection was also detected (Figure 5.3) but the situation is reversed compared to Unionoidae, i.e. the test detected slightly more sites under significant positive selection in the F genome ($-2\Delta I = 76$, p-value $\ll 10^{-5}$), with estimated proportion $p_2 = 7.8\%$, rate ratio $\omega_2 = 8.3$ and BEB inference suggesting 38 sites with $\text{Pr}(\omega_2) > 95\%$ and 85 with $\text{Pr}(\omega_2) > 75\%$ (see appendice C.3 for Supplementary Figure). In the M genomes ($-2\Delta I = 34$, p-value $\ll 10^{-5}$), we obtained an estimated proportion $p_2 = 5.3\%$, rate ratio $\omega_2 = 8.6$ and BEB inference with only 11 sites displaying $\text{Pr}(\omega_2) > 95\%$ and 41 sites displaying $\text{Pr}(\omega_2) > 75\%$ (see appendice C.4 for Supplementary Figure). The BEB estimates for the rate ratios are $\omega_2 = 6.0 \pm 6.6$ in the M genome and $\omega_2 = 18 \pm 8$ in the F genome. The large uncertainty for the estimate in the M genome gives a negative lower bound. This is no cause for concern for it is an artifact that occurs when trying to quantify the uncertainty of strongly asymmetric densities with long tails. One should use the whole PMF to get a sense of this uncertainty by inspecting Figure 5.5. Putative positively selected sites in the F genome of Veneridae are essentially localized in two genes, namely *nd4* and *nd4l* (Figure 5.6 and 5.7). This could be associated with a reorganization of mitochondrial gene order in the F genome.

Within Mytilidae, the BST did not detect any strong evidence at the significance level α_{BC} for lineages under positive selection in neither female transmitted nor male transmitted genomes (Table 5.2 and Figure 5.4). The LRT statistics for the female and male lineages are $-2\Delta I = 6.4$ (p-value = 0.011) and $-2\Delta I = 3.0$ (p-value = 0.083), respectively. In the female transmitted mitochondrial genome, the LRT did suggest weak evidence for positive selection on a small proportion $p_2 = 0.5\%$ of sites with a rate ratio ω_2

= 6.1. BEB inference is not performed on these lineages to diminish the occurrence of type I errors.

5.5 Discussion

In the present study, we used coding genes (all except *atp8*) of complete mitochondrial genome from different bivalve lineages to test if the existence of gender specific mitochondrial genome had allowed adaptive evolution. For the first time, we present data that come in support to the hypothesis that the M genome in the unusual mitochondrial inheritance system of DUI experienced positive selection.

The results obtained for the M genomes of the Unionoidae are the most notable. We found evidences that several amino acid sites of the 12 analyzed mtDNA-encoded genes involved in oxidative phosphorylation (OXPHOS) have been driven by positive selection. Those results confirm the early presumptions about DUI, i.e. this atypical system allows adaptive evolution in paternally inherited genomes (Skibinski, Gallagher and Beynon, 1994; Gemmell and Braisher, 2001; Selosse, Albert and Godelle, 2001; Zeh and Zeh, 2005). However, the exact causes for positive selection are hard to determine. It has been proposed that DUI might be a strategy to avoid sex-specific constraints associated with maternal mtDNA transmission and an opportunity for mitochondria to evolve adaptively for male function like sperm motility or longevity (Gemmell and Braisher, 2001; Selosse, Albert and Godelle, 2001; Zeh and Zeh, 2005; Everett *et al.*, 2004; Jha *et al.*, 2008). However, not all male lineages show traces of positive selection; our analyses were not able to detect any signs of positive selection in the M lineage of Mytilidae. In that context, the retained hypothesis has to explain why positive selection acts on certain male lineages and not on others. This hypothesis remains to be found. However, such a large amount of adaptive

evolution occurring on OXPHOS genes and specifically on the Unionidea lineage, might suggest that adaptive evolution of OXPHOS genes was necessary for the accomplishment of a novel function by male genome or male mitochondria. Since unionidean bivalves have several peculiar life history traits, it's plausible that in this lineage, a new or complementary positively selected function evolved.

At the genome structure level, the presence of the Mcox2 extension is the main feature that distinguishes unionoid bivalve from all other bivalves studied. The existence of a gene extension is most likely the consequence of duplication. Although the most probable outcome of gene duplication is loss of function in one of the two gene copies, duplicates could often evolve under positive selection, in which rapid changes in the protein encoded by the new gene occur to gain a new function (Roth *et al.*, 2007). Typically, the mitochondrially encoded subunit 2 (cox2) of the cytochrome c oxydase (COX) has two conserved domains; an N-terminal transmembrane helix domain and a C-terminal catalytic domain. The unionoidean bivalve M genomes possess a unique third domain (MCOX2e) at its C-terminus. This unique third domain is encoded by a 3' coding extension of the Mcox2 gene. This extension is ~185 codons in length and is likely to be the most rapidly evolving mitochondrial domain identified in animals. Structural characterization of the MCOX2e region predicted the presence of an interspecifically variable number of transmembrane helices (Chapman *et al.*, 2008), and immunohistochemistry- and immunoelectron microscopy-based analyses revealed that MCOX2e is expressed in sperm mitochondria (Chakrabarti *et al.*, 2006) and is sub-cellularly localized to both inner and outer mitochondrial membranes (Chakrabarti *et al.*, 2007). The rapid evolution, the tissular expression pattern and seasonal variation in expression profiles (Cantatore *et al.*, 1987; Chakrabarti *et al.*, 2007; Chapman *et al.*, 2008) are supporting

evidence for the hypothesis that MCOX2e has a reproductive function within unionoidean bivalves. All unionoid bivalve M genomes examined to date contain a Mcox2e region (Walker *et al.*, 2006), which is not present in other DUI-possessing bivalve lineages nor, apparently, in any other animal mitochondrial genomes (Chapman *et al.*, 2008). The exact function of Mcox2e remains unknown, and it's too early to speculate about the role of the Mcox2 extension in a putative new function of the male genome or male mitochondria, but collectively, those observations point towards it.

The presence of sites under positive selection in the F genomes of Unionoidae was slightly unexpected. The question about whether natural selection or random genetic drift is the main force in mtDNA evolution became a recurrent issue in evolutionary biology. The mitochondrial genomes from a strict maternal inheritance (SMI) system from our study (Ostreoida and Pectinoida) did not show evidence of positive selection (data not shown). In a DUI system, one could speculate that the evolutionary forces are not the same as in SMI. We speculate that, from a nucleo-mitochondrial coevolution perspective, strong positive selection acting on the M genome might have induced some adaptations in the F molecules.

Doubly uniparental inheritance has been mostly studied in the Mytilidae group. However, the LRT test suggest a very small proportion of site under positive selection (<1%) in the F genomes and BEB analysis was not able to detect any sites under positive selection in neither the M nor the F genomes. In our opinion, this is coherent with the observations that DUI seems to act differently in each bivalve lineage, which concords with observed differences in the structure and organization i.e. presence or absence of Mcox2 extension, recombination events and instability of the system. Existence of a gender-specific mtDNA probably allows a relaxation of selective constraints,

which is most likely what we observe in the M genome of Mytilidae. At the same time, this relaxation of selective constraints might have resulted in a higher level of mutation and possible novelty (i.e. Mcox2 extension) that could be retained and positively selected if they have a beneficial function, which is most likely the case from what we observed in Unionoidae.

In the Veneridae group, there are more sites under a positive selection regime in the F genome than in the M genome. In contrast to unionoid, positively selected sites are not distributed across the entire molecule but are restricted to few regions in the molecule, namely nd4, nd4l and cox2. We believe that this could result from a reorganization of mitochondrial gene order in the F genome, instead of adaptive evolution associated to DUI. NADH dehydrogenase subunits 4 and 4L genes generally overlap or are adjacent to one another in animal mt genomes and, in vertebrates nd4 and nd4l are transcribed as one bicistronic mRNA. However, in the marine clam *Venerupis philippinarum* nd4 and nd4l are separated from one another by several genes in both F and M genomes. Moreover, there is gene duplication for the cytochrome c oxidase subunit II gene (cox2) in the F genome (Mizi *et al.*, 2005; Dreyer and Steiner, 2006). Knowing that mtDNA rearrangements can have profound functional implications on gene expressions and genome replication (Gissi, Iannelli and Pesole, 2008), natural selection could act at specific sites in the region where the rearrangement occurs.

5.6 Conclusion

The study of evolutionary forces acting in the mitochondrial genomes in the DUI system has proven to draw conclusions of great general evolutionary interest. Our study demonstrates that a large amount of adaptive evolution occurred specifically on mitochondrial genes of Unionidae male lineage and to a lesser degree on the Unionidae female lineage. The branch-site test of positive selection and the Bayes empirical Bayes analyses identifies several positively selected amino acids widely distributed across all analyzed genes. The exact causes for this large positive selection remains unknown but might suggest that adaptive evolution of OXPHOS genes was necessary for the attainment of new complementary function. We suggest that the action of positive selection on the F genome of unionidean bivalves is possibly driven by the rapid evolution of M genomes. If this hypothesis is true, this could represent a manifestation of cyto-nuclear co-evolution in the DUI system. Finally, the presence of positively selected sites in veneroid F genomes could be the results of the gene order rearrangements and gene duplication. In Mytilidae, our analyses were not able to detect any signs of positive selection in neither the F nor the M lineage. The evolutionary fate of DUI appears different according to each bivalve lineage.

CHAPITRE VI

CONCLUSION GÉNÉRALE

L'objectif de cette thèse était d'étudier le rôle des contraintes sélectives sur l'évolution moléculaire des génomes mitochondriaux mâles et femelles du système de la double transmission uniparentale chez les Unionidés. L'atteinte de cet objectif a permis d'éclaircir les processus évolutifs en cause dans ce système atypique. Nos résultats font apparaître le rôle de la sélection naturelle dans l'évolution générale du génome mitochondrial. De fait, lorsqu'un système biologique est universel comme c'est le cas pour la transmission uniparentale chez les animaux, il n'y a aucun contraste; notre analyse du système ne trouve aucun point de repère. Il est alors difficile de fixer les limites et de comprendre les mécanismes en cause. La double transmission uniparentale est une des rares véritables exceptions connues à ce jour à la transmission uniparentale, et en ce sens, elle offre la possibilité d'étudier le comportement évolutif mitochondrial sous un angle différent et privilégié.

Cette thèse fait la démonstration que le système de la double transmission uniparentale est un système unique. Les différentes approches abordées dans ce travail doctoral que ce soit la revue de littérature (chapitre II), la génomique comparative (chapitre III), la phylogéographie (chapitre IV) et l'évolution moléculaire (chapitre V), apportent chacune des informations complémentaires qui se confirment et permettent de valider des hypothèses évolutives.

6.1 Principales propositions issues du chapitre II

L'analyse de portion de gènes des génomes mitochondriaux mâles et femelles d'espèces possédant la DUI ont indiqué, entre autres, que : (1) la DUI serait apparue il y a 200 millions d'années, si ce n'est pas auparavant; que (2) les génomes M et F évolue plus rapidement que chez les autres taxons animaux ne possédant pas la DUI et que (3) le génome M évolue plus rapidement que le génome F.

Les raisons de l'évolution accélérée des génomes F et M et plus particulièrement du génome M demeurent obscures. La principale hypothèse était une relaxation des contraintes sélectives attribuable à une division inégale de la charge de travail dans le système de la DUI. Contrairement à l'ADNmt typique qui doit performer pleinement dans les gonades et les tissus somatiques des deux sexes, le génome F aurait à fonctionner dans la gonade femelle et dans les tissus somatiques M et F alors que le génome M ne fonctionnerait que dans la gonade mâle.

Deux hypothèses, qui ne sont pas mutuellement exclusives, ont été proposées pour expliquer l'origine et le maintien du système de la double transmission uniparentale de l'ADNmt chez certains bivalves au cours de l'histoire évolutive : (1) la lignée mitochondriale paternelle est maintenue, car elle permet une adaptation de l'ADNmt mâle pour des fonctions spécifiques aux spermatozoïdes (par exemple la motilité, la longévité) et (2), le maintien de l'ADNmt paternel est nécessaire, car celui-ci pourrait détenir un rôle dans la détermination du sexe mâle chez les moules.

Une différence fondamentale entre la transmission maternelle stricte et la double transmission uniparentale est que cette dernière est le seul mode de

transmission de l'ADNmt animal qui permet à un génome mitochondrial d'origine paternelle de s'adapter à des fonctions mâles. Autrement dit, chez les moules, la sélection peut agir directement sur l'ADNmt paternel puisqu'il ne représente plus un cul-de-sac évolutif.

En nous offrant la possibilité d'étudier deux variants mitochondriaux fortement divergents qui coexistent et s'adaptent simultanément à un génome nucléaire, le système DUI pourrait nous permettre d'acquérir de nouvelles connaissances sur les interactions et la coadaptation des génomes nucléaires et mitochondriaux.

6.2 Principaux résultats du chapitre III

Le séquençage des ADNmt M et F chez les Unionoïdés présentait des difficultés spécifiques attribuables entre autres, au haut niveau de divergences entre les ADNmt F et M et l'inefficacité des amorces universelles. En ayant recours aux nouvelles technologies de séquençage à haut débit (Système GS-FLX 454,) il a été possible de séquencer, d'annoter et de publier pour la toute première fois des génomes mâles complets d'Unionoïdés.

Six nouveaux génomes mitochondriaux complets d'espèces d'Unionoïdés, soit le génome F et M de *Venustaconcha ellipsiformis* (Unionoida: Unionidae: Ambleminae: Lampsilini), *Pyganodon grandis* (Unionoida: Unionidae: Unioninae: Anodontini) et de *Quadrula quadrula* (Unionoida: Unionidae: Ambleminae: Quadrulini) ont été séquencés.

Parmi les groupes de bivalves possédant la DUI, les Unionoïdés se caractérisent et se distinguent par une divergence intraspécifique extrême entre les ADNmt F et M avec une moyenne de 50% pour *V. ellipsiformis*, 50% pour *I. japonensis*, 51% pour *P. grandis* et 52% pour *Q. quadrula* (distances p en acides aminés non corrigées). Du point de vue de la coévolution mitochondriale et nucléaire, comment les individus mâles peuvent tolérer une hétéroplasmie représentant autant de variabilité demeure une question non résolue, tout comme la fonction de la DUI d'ailleurs.

Les analyses phylogénétiques de 12 gènes codants de 29 bivalves et cinq génomes mitochondriaux de groupes externes indiquent une monophylie robuste des bivalves et suggèrent l'ordre suivant de ramification dans les bivalves autolamellibranches: [[Pteriomorphia + Veneroida] +Unionoida].

La reconstruction de l'état ancestral des caractères mitogénomiques (4 caractères mitogénomique analysés : gènes sur 1 brin versus 2 brins; présence/absence d'un Atp8 complet; nd4/nd4l séparés versus adjacent ou superposés; présence/absence d'un ARNtM supplémentaire) indique qu'une importante réorganisation génomique eut lieu dans les bivalves, ultérieure à la divergence de la lignée Unionoida. Pour chacun des caractères, l'état observé chez les Unionoïdés serait identique à celui de l'ancêtre des mollusques.

La position basale des Unionoïdés au sein des bivalves autolamellibranches ainsi que l'hypothétique origine unique de la DUI suggèrent que :

- la DUI serait apparue dans une lignée ancestrale des bivalves autolamellibranches;

- elle aurait été subséquemment perdue chez plusieurs descendants;
- les caractéristiques des génomes mitochondriaux observées chez les Unionoïdés s'approcheraient davantage de l'état ancestral de la DUI.

L'analyse des génomes M d'Unionoïdés a permis de mettre en évidence certaines particularités spécifiques au génome M notamment une nouvelle localisation du gène *trnH*, une inversion des gènes *ATP8-trnF* et confirme la présence d'une extension en 3' du gène de la sous-unité II de la cytochrome c oxydase. Une ou plusieurs de ces caractéristiques uniques du génome M pourrait potentiellement être liés à la transmission paternelle.

La présence de la *Mcox2e* demeure la caractéristique génomique structurale majeure qui distingue le génome M du génome F, mais qui distingue aussi le génome M des Unionoïdés de tous les autres bivalves qui possèdent la DUI et de tous les autres ADN mt des métazoaires.

6.3 Principaux résultats du chapitre IV

L'échantillonnage de trente-neuf populations de *Pyganodon* sp. sur une vaste étendue géographique et l'utilisation de deux méthodes empiriques de délimitation des espèces basée sur la séquence (General mixed Yule-coalescent method (GMYC) et Statistical parsimony network analysis) obligent à réévaluer les processus d'hybridation et de délimitation des espèces dans le genre *Pyganodon*.

Le degré d'isolation reproductif entre *P. cataracta* et *P. fragilis*, deux espèces occupant la partie est du territoire à l'étude, a été longtemps

controversé. Nos résultats supportent la distinction entre ces deux espèces et confirment le caractère récent de leur divergence.

Les analyses par séquençage moléculaire (gènes *cox1* et *cox2* des ADNmt M et F) à large échelle géographique ont révélé le caractère cryptique de la diversité génétique de la morpho-espèce *P. grandis*. Les valeurs de diversité intraspécifique de *P. grandis* mesurées sont supérieures aux valeurs de diversité interspécifiques des autres espèces de *Pyganodon*. Selon les données obtenues à partir du génome F, *P. grandis* serait en fait constitué de deux espèces distinctes. La distribution de ces dernières ne concorde avec aucun patron géographique. Toutefois, selon les données obtenues à partir du génome M, *P. grandis* serait en fait constitué de trois espèces distinctes. Ces dernières se distribuent selon un patron géographique clair correspondant à de grands bassins versants de l'Amérique du Nord.

En ce qui concerne l'extraordinaire diversité intraspécifique de *P. grandis*, plusieurs hypothèses ont été retenues. Il est possible que *P. grandis* soit en réalité un complexe d'espèces (deux espèces ou plus). Selon l'arbre phylogénétique construit à partir des séquences de l'ADN F, une possible hybridation entre *P. grandis* et une espèce de type - *P. lacustris* pourrait expliquer certains des clades supplémentaires de *P. grandis*. L'hybridation introgressive aurait possiblement été facilitée par les perturbations majeures des lacs proglaciaires lors de la dernière glaciation. Une augmentation de la diversité génétique de *P. grandis* par la fusion de groupes génétiques autrefois divergents ferait de *P. grandis*, l'espèce d'Unionoïdés la plus variée.

En ce qui concerne la situation de *P. lacustris*, les données moléculaires obtenues à partir du génome F, suggèrent que *P. lacustris* serait constituée de deux espèces distinctes se distribuant dans des régions géographiquement définies. Toutefois, un seul haplotype de *P. lacustris* aurait récemment envahi tout l'espace génotypique M conduisant à une diversité étonnamment faible, voire quasi nulle dans la lignée mitochondriale M chez cette espèce.

Une réduction marquée de diversité génétique, telle que celle observée dans l'ADNmt mâle de *P. lacustris*, peut être attribuable soit à des facteurs démographiques (par exemple un goulot d'étranglement génétique) ou à des facteurs sélectifs (par exemple, un balayage sélectif où un avantage adaptatif favorise la propagation d'un haplotype donné et efface la divergence accumulée). Cette dernière est l'hypothèse privilégiée puisque que si cette réduction de diversité génétique devait être attribuable à un goulot d'étranglement génétique, il devrait être possible de voir certaines signatures de ce goulot au niveau du génome F, ce qui n'est pas le cas. En ce sens, les facteurs sélectifs sont privilégiés.

Il ressort de ce chapitre, entre autres, le fait que les génomes M et F n'exposent pas nécessairement à des histoires phylogénétiques congruentes. Les données recueillies témoignent d'une dynamique particulière de l'évolution des deux génomes à l'intérieur des populations/espèces. L'évolution rapide du génome mâle peut permettre de mieux percevoir certains patrons géographiques récents (par exemple comme chez *P. grandis*), mais la possibilité de balayages sélectifs pouvant effacer la

divergence accumulée peut donner une fausse impression que le génome F évolue plus rapidement (par exemple *P. lacustris*).

6.4 Principaux résultats du chapitre V

Les résultats du test de sélection positive effectué en maximum de vraisemblance suggèrent que de nombreux sites d'acides aminés sont positivement sélectionnés sur le génome mâle des Unionoïdés. Cette importante pression de sélection positive agissant sur l'ensemble des 12 gènes mitochondriaux analysés pourrait suggérer qu'une évolution adaptative des gènes de la phosphorylation oxydative était nécessaire à l'émergence d'une nouvelle fonction complémentaire pour les mitochondries mâles ou pour le génome mâle chez cette lignée de bivalves. Cette hypothétique nouvelle fonction complémentaire reste à identifier.

Quelques sites seraient également positivement sélectionnés dans l'ADNmt F des Unionoïdés et témoignent possiblement de la coévolution cytonucléaire. La forte pression sélective agissant sur le génome M pourrait avoir entraîné au passage le génome F à s'adapter aux ajustements au niveau du génome nucléaire.

Le test n'a pas mis en évidence de sélection positive dans les ADNmt M ni F de Mytilidés et confirment que les forces évolutives n'agissent pas de manière identique dans les différents groupes de bivalves possédant la DUI.

Les résultats du test révèlent que des sites d'acides aminés seraient également positivement sélectionnés chez les Vénéridés. Toutefois, la

concentration de ces sites sur certains gènes de l'ADNmt F pourrait témoigner de l'action de la sélection positive suite à une réorganisation structurale du génome.

Le système de la double transmission uniparentale a été découvert et nommé dans les années 1990. Il s'agit d'une des rares véritables exceptions connues à la stricte transmission uniparentale chez les animaux. Comprendre et élucider les forces évolutives en cause dans un système de transmission mitochondrial atypique étaient un objectif audacieux. Au commencement de cette thèse, les informations disponibles sur les Unionoïdés étaient éparses et celles chez les taxons marins pointaient vers différentes conclusions. En abordant l'objectif général de ce projet de thèse en plusieurs étapes successives d'envergure soit (1) la caractérisation et la comparaison de génomes complets, (2) l'analyse de la variabilité de gènes des ADNmt M et F entre populations et entre espèces proches, et (3) l'analyse et la comparaison de l'évolution moléculaire des ADNmt M et F entre eux et avec des ADN mitochondriaux d'espèces de bivalves dont la transmission se fait de façon maternelle standard, il fut possible de commencer à dégager une image claire des processus évolutifs en cause.

6.5 Généralisation et réflexions sur l'origine de la DUI

La nature même des eucaryotes est définie par la présence d'organites intracellulaires, les mitochondries et les plastes, lesquels possèdent leur propre génome. Les processus cellulaires vitaux dépendent de ces organites et de leurs coordinations fonctionnelles avec le noyau. Comprendre comment et pourquoi les génomes des organites se sont maintenus au cours de l'évolution représente une question fondamentale en biologie évolutive. En ce

sens, l'existence même du système de la DUI, soulève de nombreuses questions.

L'ensemble des résultats de cette thèse et les réflexions finales qui en découlent amènent à un questionnement sur la généralisation du système de la DUI. De fait, il est possible d'envisager et d'analyser le système de la DUI comme un génome dupliqué. Les duplications peuvent survenir au niveau de gènes spécifiques, de fragments de chromosomes ou peuvent être globales, soit une duplication du génome en entier.

Les duplications génomiques sont au cœur du processus d'évolution dans les trois domaines du vivant, soit les bactéries, les archaea, et les eucaryotes (Zhang, 2003). La duplication de gènes et de génomes est présentée comme un élément moteur de l'évolution facilitant l'apparition de nouvelles fonctions (Bassham, Canestro and Postlethwait, 2008), l'adaptation à de nouveaux environnements (Hittinger and Carroll, 2007) et la spéciation (Scannell *et al.*, 2006).

Les mécanismes qui génèrent la duplication de gènes ou de portion de gènes sont divers, ils incluent la rétrotransposition, les enjambements inégaux, les échanges ectopiques et le transfert horizontal de gènes. Les duplications totales de génome nucléaire (polyploïdie) sont engendrées par autopolyploïdie ou allopolyploïdie. Bien que la duplication de génomes complets de mitochondries ou de plastes ne figure généralement pas dans la littérature associée à duplication de gènes et de génomes, il est facile d'imaginer plusieurs mécanismes via lesquels une duplication pourrait survenir tel qu'une perturbation du processus de réplication, la recombinaison

ou, en fait, toutes perturbations des mécanismes responsables du maintien de l'homoplasme.

Une fois dupliqué, le gène ou le génome peut avoir différentes destinées évolutives (Ohno, 1970; Force *et al.*, 1999; Lynch and Conery, 2000; Moore and Purugganan, 2003). De façon générale, la grande majorité des gènes ou des génomes dupliqués seront éliminés très rapidement par la dérive et ou la sélection naturelle. Quelques-uns remplaceront leurs copies parentales et très peu expérimenteront des changements fonctionnels qui assurent leur préservation à long terme.

Spécifiquement, les voies possibles d'évolution des gènes dupliqués une fois fixés sont (Figure 6.1) :

- la non fonctionnalisation (ou pseudogénisation) : une des copies accumule des mutations et dégénère en passant par l'état de pseudogène.
- la redondance : la copie dupliquée conserve la même fonction que la copie originelle ;
- la subfonctionnalisation : la fonction ancestrale est partitionnée et chaque copie assure une partie du partage des fonctions ancestrales. La subfonctionnalisation peut également correspondre à la division de l'expression génique après une duplication c'est-à-dire que les deux copies seront exprimées dans des conditions spatiales ou temporelles différentes, soit exprimées dans différents tissus ou exprimées à différents stades du développement. Dans les deux cas, il y a une complémentarité fonctionnelle entre les deux copies et elles seront conservées dans le génome.

– la néofonctionnalisation : les deux copies divergent. Une des copies va conserver la fonction originelle alors que l'autre va acquérir une nouvelle fonction.

Enfin, puisque ni le modèle de la subfonctionnalisation ni celui de la néofonctionnalisation arrivaient à bien expliquer les données expérimentales, He et Zhang (2005) proposèrent un modèle évolutif intermédiaire soit la subnéofonctionnalisation. Selon ces auteurs, il y aurait d'abord un rapide événement de subfonctionnalisation. Par la suite, les deux copies deviendraient essentielles, maintiendraient leur patron d'expression originale et seraient donc conservées. Une fois que le ou les gènes sont établis dans le génome, ils peuvent conserver leur fonction ou évoluer et se spécialiser à nouveau (c'est-à-dire subir une néofonctionnalisation). En conséquence, la subfonctionnalisation semble être un processus rapide, tandis que les néofonctionnalisations exigent plus de temps et se poursuivraient même longtemps après la duplication (He and Zhang, 2006).

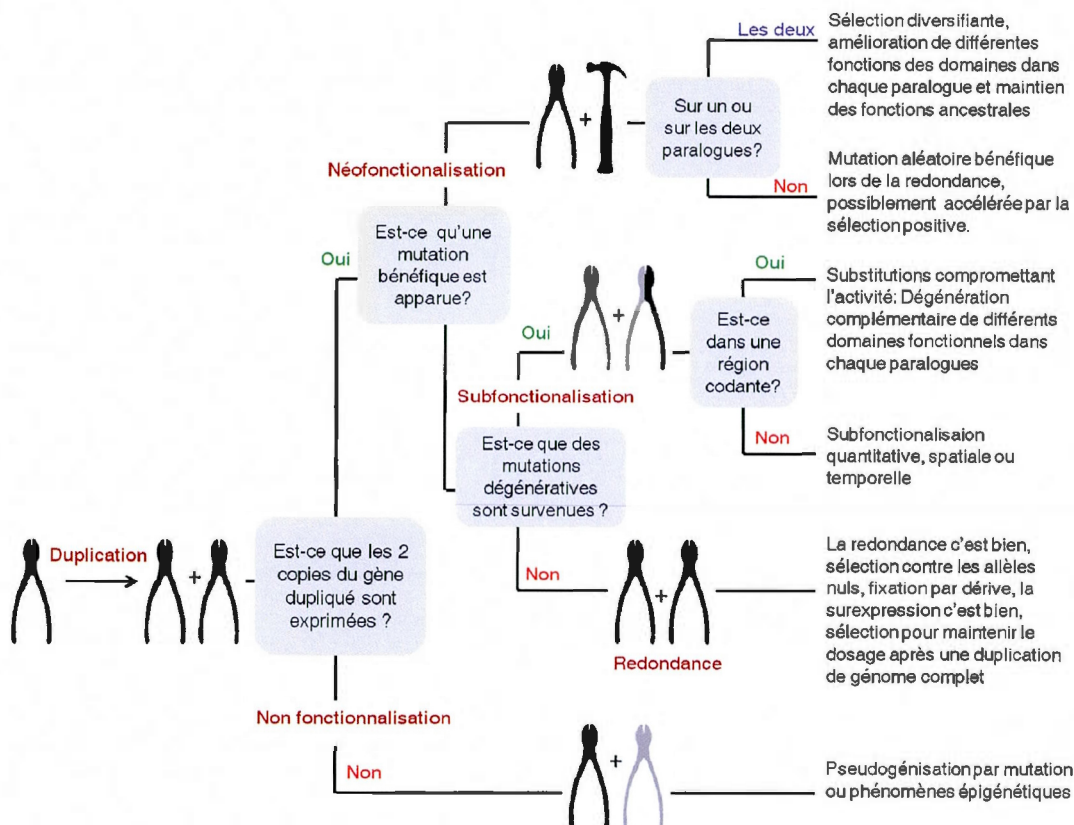


Figure 6.1 Les destinées évolutives possibles des gènes dupliqués incluent la néofonctionnalisation, la subfonctionnalisation, la redondance et la non fonctionnalisation. Ces voies d'évolution sont la conséquence de la sélection naturelle, des mutations et de processus épigénétiques qui agissent sur les régions de régulation et sur les régions codantes des gènes dupliqués. Dans cette figure, une protéine est représentée par une paire de pinces, chacune des branches de la paire représente une sous-fonction distincte de l'outil. La néofonctionnalisation transforme une des copies en un nouvel outil (un marteau) avec une nouvelle fonction. La subfonctionnalisation dégrade une des branches complémentaires de la protéine dupliquée (grise pâle) rendant l'activité combinée des deux paralogues nécessaire pour accomplir la fonction ancestrale. La non fonctionnalisation dégrade complètement une des protéines dupliquées. Inspirée de Evans (2008).

En considérant ce cadre théorique, les questions soulevées par le système de la DUI sont : est-ce possible que les voies d'évolution que nous observons dans le système de la DUI s'apparentent au dessin évolutif d'un génome dupliqué ? Les détails du scénario d'origine de la DUI ne tiendraient-ils dans une simple duplication de génome ? Le génome mitochondrial mâle des Unionoïdés ne représenterait-il pas un exemple éloquent de subfonctionnalisation ? La *cox2* extension serait-elle un exemple de néofonctionnalisation ?

Les résultats du chapitre III suggèrent une origine unique de la DUI. Nous pourrions spéculer sur l'idée qu'une duplication d'un groupe de gènes nucléaires régissant le maintien, la réplication et les fonctions du génome mitochondrial auraient permis l'existence de deux variants du génome mitochondrial (Figure 6.2). Ces deux copies mitochondriales, probablement sans aucune autre fonction spécifique à l'époque (que celles d'assurer les fonctions mitochondriales déjà établies), auraient coexisté dans une lignée ancestrale des bivalves autolamellibranches et auraient été subséquentement perdues chez plusieurs descendants. Se fixant seulement chez certains groupes, une subfonctionnalisation aurait résulté en un variant mitochondrial spécifique au tissu reproducteur mâle (subfonctionnalisation spatiale) (Figure 6.2). Cette voie évolutive pourrait avoir été celle empruntée par les Vénéridés et les Mytilidés selon les résultats du chapitre V.

Chez les Unionidés, une néofonctionnalisation subséquente aurait permis l'apparition d'une nouvelle fonction, les résultats des chapitres III et IV laissent croire que celle-ci pourrait être associée avec l'extension fonctionnelle du gène de la cytochrome c oxydase. Une importante sélection positive (résultat du chapitre V) accélérerait l'évolution de l'ensemble des gènes

du génome mâle. Pour l'instant la fonction exacte de cette extension demeure inconnue. Toutefois, cette extension pourrait être, telle qu'actuellement proposée, une protéine de la reproduction. La stabilisation d'un gène dupliqué dépend de la pression de sélection pour la fonction assurée. Les gènes impliqués dans les fonctions immunitaires et reproductives qui varient entre les espèces montrent plus fréquemment des signatures de sélection positive (Swanson and Vacquier, 2002).

En appui à ce scénario, soulignons qu'au cours de l'évolution des mollusques bivalves, des changements significatifs dans la structure et le nombre de chromosomes sont survenus (Wang, Xu and Guo, 2005; Leitao and Chaves, 2008). Il est même suggéré qu'au moins une duplication de génome entier se soit produite dans l'évolution des bivalves (Wang and Guo, 2004). Au sein de ce groupe, on observe une grande diversité de caryotypes entre espèces. Par exemple, les espèces d'huîtres (Ostréidé) ont 20 chromosomes ($2n=20$), les Mytilidés en ont de 28 à 32 ($2n$) selon les espèces, et les Vénéridés, les Unionidés et les Donacidés possèdent 38 chromosomes ($2n$) (Wang, Xu and Guo, 2005; Leitao and Chaves, 2008). Il est par conséquent vraisemblable qu'une duplication d'une partie de chromosome, une duplication entière d'un chromosome ou une duplication d'un génome nucléaire entier soit à l'origine de la duplication de l'ensemble des gènes nucléaires nécessaires au maintien et à la régulation d'un deuxième variant mitochondrial (Figure 6.2).

Initialement deux hypothèses, qui ne sont pas mutuellement exclusives, avaient été proposées pour expliquer l'origine et le maintien du système de la DUI soit que la lignée mitochondriale paternelle ait été maintenue puisqu'elle permettrait une adaptation de l'ADNmt mâle pour des fonctions spécifiques

aux spermatozoïdes et/ou que le maintien de l'ADNmt paternel soit nécessaire, car celui-ci détiendrait un rôle prépondérant dans la détermination du sexe mâle chez les moules. Toutefois, ces hypothèses ne proposaient aucun mécanisme par lequel le système de la DUI aurait pu émerger. À la lumière des différents chapitres de cette thèse et des réflexions qui en découlent, il est possible de faire une relecture du système de la DUI et de ces implications. Tout semble indiquer que la lignée paternelle puisse effectivement donner naissance à des nouveautés évolutives, toutefois cela n'explique en rien l'origine ou le maintien du système de la DUI. L'extension de la *cox2* (*Mcox2ext*) pourrait expliquer l'absence d'événement de masculinisation chez les bivalves unionoïdes et la forte sélection positive observée chez les bivalves unionïdes. En ce sens, l'apparition d'une nouvelle extension fonctionnelle impliquée dans la reproduction contribuerait à la stabilité du système chez ce groupe.

En ce qui concerne le rôle du génome mitochondrial mâle dans la détermination du sexe, certaines informations et réflexions supplémentaires permettent également de nuancer cette proposition. Malgré certains progrès (Yusa, 2007), le mécanisme de détermination du sexe est encore inconnu chez les bivalves. Les bivalves sont généralement gonochoriques, mais certains individus, certaines populations et certaines espèces présentent un hermaphrodisme. Chez *M. galloprovincialis*, tous les individus triploïdes sont mâles (Kiyomoto *et al.*, 1996). Un sex-ratio fortement biaisé vers les mâles est également confirmé lors de la production de triploïdes chez *M. edulis* (Brake, Davidson and Davis, 2004). Ces résultats permettent d'envisager que le déterminisme du sexe soit génétique chez ce groupe et serait de type Z:O (de chromosome Z unique) ou Z:W (les femelles ont deux chromosomes de différentes sortes). Chez les bivalves unionidés, le déterminisme du sexe est

complètement inconnu et ce, y compris chez les espèces où les sexes sont séparés de même que chez les quelques rares espèces qui utilisent l'hermaphrodisme comme mode de reproduction principal (Kat, 1983). L'analyse des sex-ratios et de l'hermaphrodisme chez *Margaritifera margaritifera* avait amené Bauer (1987) à suggérer que la détermination du sexe devait être un simple mécanisme génétique d'hétérogamétie.

Dans une étude convaincante, basée sur des croisements hybrides, à partir de moules triploïdes et des observations des mitochondries du sperme dans les œufs fécondés, Kenchington *et al.* (2009) montre clairement que la masculinité et la présence du génome mitochondrial paternel peuvent être dissociées chez les Mytilidés. Chez les Unionidés, Breton *et al.* (2011) furent incapables de trouver la présence du génome M chez des espèces d'unionidés hermaphrodites. Ces nouveaux développements ne permettent pas encore de trancher la question définitivement. Il est possible que la situation ne soit pas exactement la même chez les Mytilidés et chez les Unionidés. Cela tend tout de même à écarter l'hypothèse d'un rôle prépondérant dans la détermination du sexe pour expliquer l'origine et le maintien de la DUI. Les lacunes de nos connaissances en ce qui concerne le mécanisme de détermination du sexe chez les bivalves demeurent pour le moment encore trop importantes pour élucider la relation entre l'ADNmt mâle et la masculinité. En fait, il ne fait nul doute que le génome mâle soit, d'une quelconque façon, lié aux gènes responsables de la détermination du sexe, mais la relation semble plus associative que causale.

Ces propositions concernant la fonction du système de la DUI n'en sont encore qu'au stade de réflexion et n'ont pas encore toutes été formulées sous forme d'hypothèses. Les discussions et conclusions de chacun des

chapitres de cette thèse ont toutefois contribué au développement de mes propres réflexions. Alors qu'au commencement de cette thèse, le système de la DUI apparaissait principalement comme une curiosité ou une pure exception qui pouvait être exploitée pour tenter de valider différentes questions évolutives. Cela en faisait certes, un système modèle intéressant, mais difficile à généraliser. Le système de la DUI, présenté comme un exemple de duplication de génome mitochondrial et de néofonctionnalisation se révèle d'un intérêt scientifique encore plus grand. À l'intérieur de ce cadre conceptuel, l'origine du système de la DUI, le maintien de ce système de transmission atypique, l'action de la sélection sur les différents génomes mitochondriaux mâles et femelles, les différences entre les groupes de bivalves prennent un sens cohérent pouvant être généralisé.

Si les travaux de la présente thèse devaient être revisités ou poursuivis, il serait assurément intéressant de le faire sous cet angle et de vérifier l'exactitude de ces réflexions et spéculations.

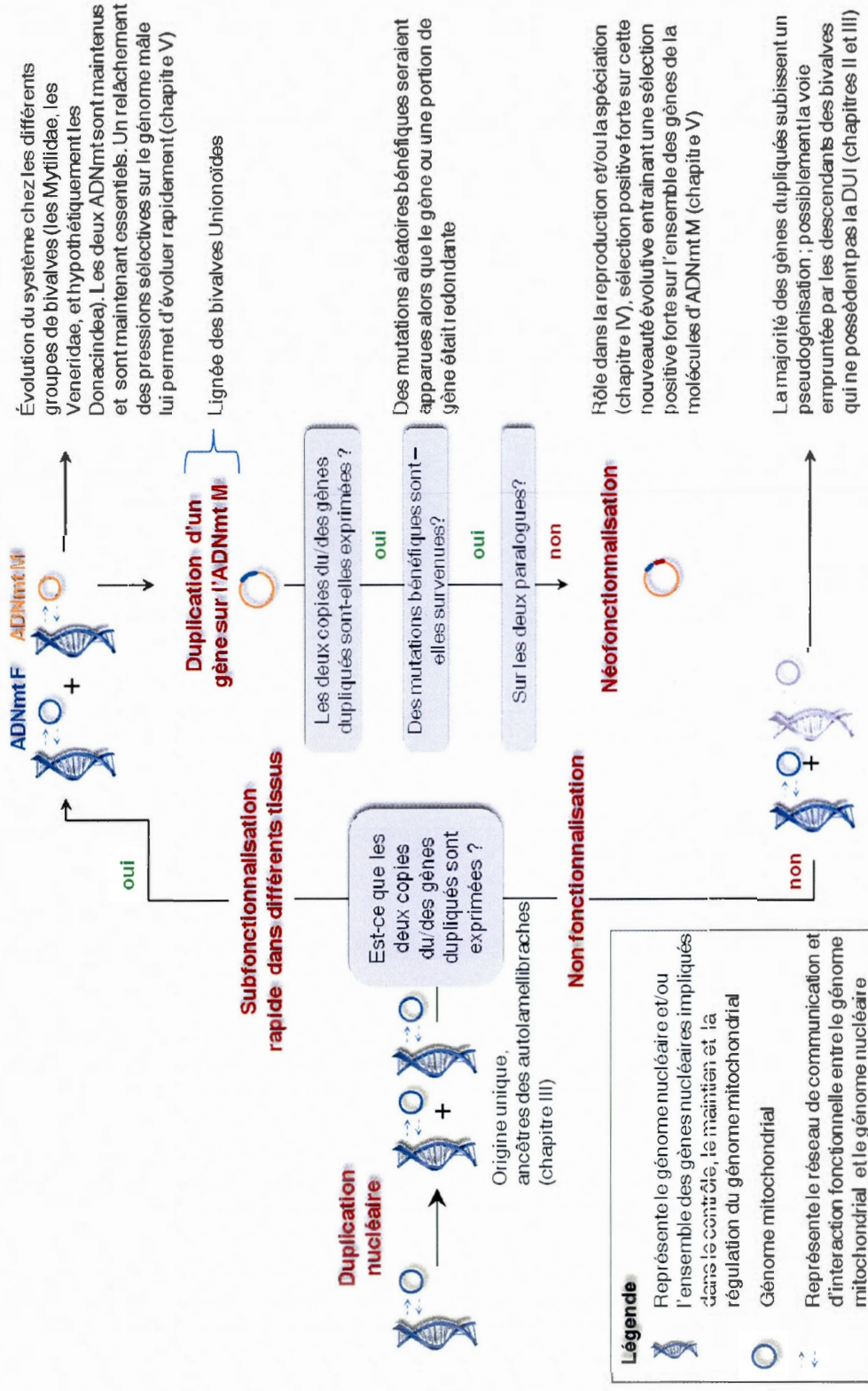


Figure 6.2 Diagramme conceptuel de l'origine de la DUI et de son évolution selon les modèles de duplication génique et génomique connus.

6.6 Contributions et limites du système de la DUI pour l'étude de la coévolution des génomes mitochondriaux et nucléaire

Au commencement de cette thèse, il a été proposé que la DUI soit un système modèle pour l'étude de la coévolution des génomes mitochondriaux et nucléaires. Les principaux arguments avancés étaient qu'en nous offrant la possibilité d'étudier deux variants mitochondriaux qui coexistent et s'adaptent simultanément à un génome nucléaire, le système DUI pourrait nous permettre d'acquérir de nouvelles connaissances sur les interactions et la coévolution des génomes nucléaires et mitochondriaux. Les taux de divergences des génomes F et M des moules d'eau douce, affichent par ailleurs les plus hautes valeurs intraindividuelles jamais rapportées. Les premières estimations basées sur les variations dans la *cox1* nous laissaient croire que la divergence était d'environ 30%. Cela était déjà plus élevé encore que dans les différents croisements réalisés entre les populations ou entre les espèces dans les modèles classiques d'étude de coévolution. Toutefois, les résultats du chapitre III basés sur l'étude de la variation de l'ensemble des gènes mitochondriaux des génomes M et F démontrent que la divergence en acides aminés entre les variants mitochondriaux M et F se situent plutôt autour de 50%. L'intérêt du système de la DUI dans l'étude de la coévolution des génomes mitochondriaux et nucléaires venait du fait que contrairement aux croisements ou aux injections de cytoplasme entre les populations ou entre les espèces réalisés en laboratoire, les deux lignées mitochondriales, même si elles sont fortement divergentes, évoluent depuis des centaines de millions d'années avec un génome nucléaire commun.

Bien qu'aucune preuve expérimentale ne supporte cette argumentation, les résultats et l'avancement de la réflexion sur le système de la DUI

semblent indiquer que le système de la DUI pourrait constituer en lui-même une démonstration qu'un variant mitochondrial doit nécessairement coévoluer avec un ensemble de gènes nucléaires spécifiques. L'hypothèse est que cette pression de coévolution forte ne peut permettre l'existence d'un deuxième variant divergeant puisque toute modification du dialogue entre les mitochondries et le noyau risquerait de provoquer l'apparition de phénotypes indésirables et de la compétition intermitochondriale de même qu'une perturbation des interactions fonctionnelles entre les protéines codées par l'ADNnu et l'ADNmt. C'est donc dire, que pour qu'un deuxième variant mitochondrial puisse exister, un ensemble de gènes nucléaires ont dû être dupliqués. Dans le chapitre II, la possibilité que certaines isoformes nucléaires existent avait déjà été évoquée (par exemple une isoforme spécifique aux gonades mâles de cytochrome c ou de la sous-unité VIb de la cytochrome c oxydase telle qu'observée chez certains mammifères (Hüttemann, Jaradat and Grossman, 2003). Toutefois, il semble plus plausible qu'un grand groupe de gènes nucléaires entrent spécifiquement en relation et en interaction avec le génome mitochondrial mâle. Ainsi, pour reprendre le même vocabulaire que précédemment, dans le système de la DUI, il y a possiblement non pas deux variants mitochondriaux pour un environnement nucléaire commun, mais deux variants mitochondriaux dans deux contextes nucléaires complètement ou partiellement distincts. Ceci est une hypothèse à vérifier et une avenue de recherche prometteuse.

Néanmoins, si cette dernière hypothèse se révélait exacte, cela ferait tout de même du système de la DUI un modèle intéressant et unique pour l'étude de la coévolution des génomes mitochondriaux et nucléaires. Il pourrait permettre l'identification de gènes spécifiques ou de groupes de gènes spécifiques nécessaire au maintien d'un deuxième variant. Il pourrait

permettre d'identifier et de caractériser les gènes nucléaires qui coévoluent obligatoirement avec le génome mitochondrial ; ces derniers ne se restreignant probablement pas seulement aux composantes structurales de du système de transport des électrons, mais ils incluent vraisemblablement aussi les gènes nucléaires responsables de la réplication, de la transcription et de la traduction de l'ADN mitochondrial de même que des gènes responsables de sa distribution dans différents tissus. L'identification de ces gènes pourrait constituer une avancée significative dans nos connaissances en biologie mitochondriale.

6.7 Signification des résultats pour la génétique mitochondriale et la biologie évolutive

L'évolution de l'ADN mitochondrial est un sujet central dans de nombreux domaines de recherche que ce soit paléogéographie humaine, dans l'étude des maladies génétiques, dans l'étude de la mort cellulaire programmée (apoptose), dans le domaine de la fertilité, du vieillissement, de la bioénergétique, de l'origine des sexes et l'évolution des cellules eucaryotes.

De façon générale, la présente thèse contribue à l'avancement des connaissances de base en génétique et en biologie mitochondriale. Par exemple, le chapitre II publié dans la revue *Trends in Genetics* fut cité par de nombreux auteurs œuvrant sur des projets de recherche rapprochés ou connexes au système de la DUI, mais également par des auteurs travaillant dans les domaines plus éloignés tels que la médecine légale de la faune (Alacs *et al.*, 2010) ou dans le clonage d'animaux de ferme (Burgstaller *et al.*, 2007).

Plus spécifiquement, les résultats obtenus dans le cadre de ce projet de doctorat occupent une place significative en génomique comparative et réitèrent l'importance de l'analyse comparée de génomes mitochondriaux complets, y compris entre espèces proches. De fait, l'étude de l'évolution des génomes peut être abordée par différentes stratégies. Les génomes ne sont pas des entités statiques, ce sont des territoires dynamiques qui évoluent soit par mutations locales (au niveau de la séquence) ou par réorganisations structurales (au niveau du génome). Pour des considérations pratiques, les premières études de l'évolution des génomes se concentrèrent principalement sur les divergences nucléotidiques de gènes spécifiques. Cependant avec la prolifération de séquences de génomes complets, nous assistons à l'émergence d'une vision plus globale de l'évolution génomique et en particulier de la génomique mitochondriale. Tout comme les séquences nucléotidiques qui mutent constamment à l'intérieur des populations, les modifications au niveau des régions chromosomiques entraînent fréquemment des insertions, des déplacements, des suppressions ou des duplications. Une variation extraordinaire a été observée en ce qui a trait à la taille du génome, au contenu génique et à l'arrangement du génome mitochondrial chez les bivalves (Wang, Xu and Guo, 2005; Leitao and Chaves, 2008). Les résultats du chapitre III fournissent une fondation solide pour la comparaison de génomes mitochondriaux chez ce groupe, où les événements macrogénomiques tels que les événements de duplication en tandem et les réarrangements (ordre des gènes et distribution entre les brins) peuvent être étudiés afin de mieux comprendre l'évolution du génome mitochondrial. À partir des réarrangements observés au sein de la famille des Vénéridés, il a été proposé que l'échange de blocs entre gènes adjacents puisse être un phénomène répandu dans l'évolution des génomes mitochondriaux chez ces espèces (Xu, Wu and Yu, 2010). Qui plus est, les

résultats du chapitre V démontrent que ces réarrangements pourraient être positivement sélectionnés. Ces résultats démontrent, pour la toute première fois l'action de la sélection positive sur un réarrangement mitochondrial.

Ce dernier résultat revêt une importance particulière puisque le séquençage et la caractérisation de génomes mitochondriaux complets ne cessent de révéler la variabilité de l'organisation de ce dernier, et ce, même chez les groupes d'organismes où, autrefois, l'ordre des gènes était considéré comme relativement stable. Par exemple, des déviations de l'ordre classique ont été trouvées chez de nombreux chordés et des lignées de vertébrés incluant les amphioxus (Boore, Daehler and Brown, 1999), les lamproies (Lee and Kocher, 1995), les poissons osseux (Inoue *et al.*, 2001), les amphibiens (Macey *et al.*, 1997; Mueller and Boore, 2005; Kurabayashi *et al.*, 2008), les lézards (Amer and Kumazawa, 2007), les oiseaux (Mindell, Sorenson and Dimcheff, 1998) et les mammifères (Janke *et al.*, 1994).

Les résultats de la présente thèse ont également un impact dans le domaine de la phylogéographie. Rappelons que l'ADN mitochondrial des métazoaires est un marqueur génétique de choix dans la reconstruction des relations phylogénétiques, et ce, à plusieurs niveaux taxonomiques, des populations jusqu'au phylum, et il a également été largement utilisé pour résoudre des situations taxonomiques complexes (Gissi, Iannelli and Pesole, 2008). Compte tenu de la complexité du système de la DUI et des forces évolutives en cause, cette thèse démontre clairement qu'une caractérisation de la biodiversité des espèces de moules d'eau douce basée seulement sur une analyse génétique de marqueurs mitochondriaux issus du génome F (par exemple, l'utilisation du code-barres génétique (de l'anglais DNA Barcoding) c.-à-d. *cox1*) serait incomplète. Par exemple, si les espèces de *Pyganodon*

avaient été analysées avec le code-barre génétique seulement, la grande variabilité de *P. grandis* aurait passé inaperçue (selon les analyses du génome F *P. grandis* serait deux espèces et selon les analyses du génome M *P. grandis* serait trois espèces).

Les marqueurs mitochondriaux ont été et continuent d'être des marqueurs largement utilisés pour décrire des patrons phylogéographiques et phylogénétiques. Dans l'histoire récente de la biologie moléculaire, l'ADN mitochondrial en tant que marqueur a connue deux périodes de popularité. D'abord, alors que le concept de marqueur moléculaire était en plein essor, plus de huit études phylogéographiques sur dix utilisaient l'ADN mitochondrial pour retracer l'histoire des populations et il en fut ainsi, jusque dans les années 2000 (Hickerson *et al.*, 2010). Dans ce contexte, la plupart des études assumaient que les variations de l'ADNmt étaient sélectivement neutres ou quasi neutres et que la variation de l'ADNmt reflétait pleinement une histoire des populations non biaisée par la sélection ou les flux de gènes. Alors que la popularité de l'ADNmt comme marqueur moléculaire grandissait, sa neutralité et les limites de son utilisation commencèrent à être sérieusement remises en question par les biologistes de l'évolution (Avise, 1994). Pourtant, à la même période et malgré cette contestation, l'ADNmt entama sa seconde ère de popularité avec l'avènement de méthodes d'identification taxonomiques rapides comme le code-barres génétique (Hebert *et al.*, 2003).

Le débat sur la pertinence de l'utilisation des marqueurs mitochondriaux en phylogénie et phylogéographie est toujours d'actualité (Hajibabaei *et al.*, 2007; Zink and Barrowclough, 2008; Barrowclough and Zink, 2009; Edwards and Bensch, 2009; Spooner, 2009). L'objectif de cette thèse n'est pas de

prendre position dans ce débat. Néanmoins, collectivement, les résultats du chapitre IV et V pourront alimenter les réflexions sur le sujet. D'une part, les résultats du chapitre IV démontrent la mosaïque de forces évolutives et le caractère complexe de leur interaction, de même que les difficultés associées à la délimitation des espèces chez les bivalves unionidés. D'autre part, les résultats du chapitre V montrent de façon éloquente que la sélection positive peut agir sur le génome mitochondrial, spécialement si ceux-ci sont sujets à des réorganisations génomiques. Évidemment, le système de la DUI est un système particulier et les forces qui agissent sur lui le sont probablement tout autant. Il serait inadéquat d'extrapoler trop rapidement ces résultats aux organismes ne possédant pas la DUI.

6.8 Signification des résultats pour la biologie et l'écologie des Unionoïdés

Bien que le sujet principal de cette thèse soit l'évolution du système de la DUI, les bivalves unionoïdes en étaient le principal modèle animal et les travaux et les mesures des différents chapitres furent principalement réalisés chez ce groupe d'organismes. Conséquemment, cette thèse comporte des données importantes concernant la biologie et l'écologie des Unionoïdés.

Les moules d'eau douce (Ordre des Unionoida) incluent six familles² : les Unionidae (Eurasie, Afrique, Inde et Amérique du Nord) ; les Margaritiferidae (Eurasie, Amérique du Nord) ; les Hyriidae (Australie et Amérique du Sud) ;

² La classification des moules d'eau douce a été changée en octobre 2011. Auparavant divisées en deux superfamilles (Etherioidae et Unionoidea) selon Graf et Cummings (2007), elles sont maintenant divisées en six familles suivant les travaux de Graf et Cummings (2006), Bieler, Carter et Coan (2012) et Whelan, Geneva et Graf (2011).

les Etheriidae (Afrique, Amérique du Sud et Inde) ; les Mycetopodidae (Amérique du Sud) et les Iridinidea (Afrique) (Walker *et al.*, 2001; Graf et Cummings, 2006; Bieler, Carter et Coan, 2010; Whelan, Geneva et Graf, 2011).

La plus grande diversité de moules d'eau douce se retrouve en Amérique du Nord, avec près de 300 espèces (Williams *et al.*, 1993; Turgeon *et al.*, 1998). Par contre, les Unionidés seraient le groupe animal le plus menacé au monde (Lydeard *et al.*, 2004) ou près de 70% des espèces indigènes sont vulnérables, c'est-à-dire soit en danger, menacées d'extinction ou ont un statut précaire. Plusieurs facteurs peuvent être mis en cause pour expliquer le déclin accéléré des moules d'eau douce indigènes observés en Amérique du Nord. La dégradation des habitats, la détérioration de la qualité des eaux et la récente introduction de la moule zébrée (*Dreissena polymorpha*), une espèce exotique invasive, seraient parmi les principaux facteurs du déclin (Nalepa, 1994; Schloesser and Nalepa, 1994).

Les moules d'eau douce ont un rôle trophique crucial dans nos plans d'eau, mais puisqu'aucune raison socio-économique ne motive l'acquisition de connaissances (les moules d'eau douce sont non comestibles et ne font l'objet d'aucune pêche ni d'aucun élevage), elles demeurent sous-étudiées et encore méconnues comparativement à d'autres espèces dulcicoles.

La conservation implique la protection et le maintien, par un ensemble de mesures d'intensité variable de la diversité génétique, des espèces, des écosystèmes et des phénomènes évolutifs auxquels ils sont soumis. Un plan de conservation efficace des Unionidés implique de protéger à la fois les espèces, mais aussi les populations. En ce sens, il devient urgent d'acquérir

des connaissances de base sur la distribution des différentes espèces, mais également une compréhension approfondie de la variation génétique entre et à l'intérieur des populations est essentielle.

Les résultats présentés dans le chapitre IV font partie des rares données à grande échelle que nous possédons sur la phylogénie et la distribution des moules d'eau douce dans le nord-est de l'Amérique du Nord. En fait, ce sont les seules données disponibles et aussi détaillées pour le territoire et pour le groupe d'espèces à l'étude, les *Pyganodon*.

L'identification des espèces d'Unionidés se fait traditionnellement à partir des coquilles, mais la très grande plasticité phénotypique peut rendre l'identification de certains groupes très difficile. Les espèces du genre *Pyganodon* analysées dans le chapitre IV sont un groupe où l'identification par les traits coquilliers demeure ambiguë (Cyr *et al.*, 2007). Les données et analyses de délimitation des espèces présentées dans le chapitre IV révèlent une grande diversité cryptique au sein de ce complexe d'espèces et soulignent l'importance de l'utilisation de marqueurs moléculaires. Si un plan de conservation devait être établi, les données du chapitre IV pourraient permettre aux gestionnaires de mieux cibler les efforts sur certaines populations présentant des particularités génétiques. Et bien que le chapitre 3 mette en évidence une grande variabilité génétique au sein des espèces de *Pyganodon*, cette dernière est issue d'un échantillonnage limité et ne prétend pas représenter la diversité réelle. En fait, la diversité réelle pourrait bien être supérieure à celle mesurée dans les données du chapitre IV.

De nombreux facteurs, propres aux moules d'eau douce indigènes, peuvent profondément influencer la diversité génétique et les variations

géographiques observées que ce soit la nature insulaire des habitats dulcicoles (Strayer, 2006), leurs capacités de dispersion limitées, le haut degré de variation et d'endémisme, leur association avec de grands systèmes de drainage qui ont une longue histoire géologique tout comme certains de leurs traits d'histoire de vie complexes et uniques. Les moules d'eau douce indigènes sont ovovivipares et leur reproduction inclut une phase larvaire parasitaire sur un poisson-hôte.

D'un point de vue écologique, les raisons qui justifient un effort de conservation des moules d'eau douces indigènes sont nombreuses. Soulignons entre autres que (1) en tant qu'organisme filtreur, elles influencent la qualité de l'eau, (2) elles influencent la dynamique des nutriments dans les écosystèmes, en filtrant les nutriments inorganiques qui ensuite, par la biodéposition, deviennent assimilables pour les plantes et animaux, (3) elles contrôlent la densité de phytoplancton ; (4) elles jouent un rôle dans le ralentissement de l'eutrophisation des lacs ; (5) elles sont une source de nourriture pour les rats musqués, les loutres, les visons, les ratons-laveurs et certaines espèces de poissons, (6) leurs coquilles influencent les caractéristiques hydrologiques du fond des plans d'eau, elles modifient l'habitat physique, et influencent la dynamique des sédiments et des courants (Paquet *et al.*, 2004).

Toutefois, si l'étude de la variation intra et interspécifiques des génomes F et M du chapitre IV amène de précieuses connaissances sur la délimitation des espèces, les relations évolutives, et les niveaux de variation dans les populations. Les chapitres II, III et V mettent en lumière que la conservation et l'étude de la génétique des Unionoïdés revêtent une importance toute particulière du fait qu'elles possèdent la double transmission uniparentale. Le

système de la DUI est un modèle inestimable et jusqu'à maintenant encore unique. Cet argument, rarement évoqué, devrait pourtant constituer une motivation majeure à l'étude et la conservation de ces espèces. Qui plus est, les chapitres III et V démontrent que le système de la DUI chez les Unionoïdés possède des caractéristiques uniques, absentes des autres groupes de bivalves possédant la DUI.

Les chapitres de cette thèse établissent collectivement la grande valeur scientifique des Unionidés et constituent un appui supplémentaire à leur étude et leur conservation.

6.9 Perspectives de recherche

Dans le cadre de ce projet de doctorat, de nombreuses avancées dans notre compréhension de la structure, du fonctionnement, de l'organisation et de l'évolution de la DUI ont été réalisées. Par contre, de nombreuses données biologiques de base manquent encore sur ce système. Par exemple, une meilleure connaissance de la distribution taxonomique du système de la DUI apparaît comme un préalable à l'avancement de l'étude de l'évolution de la DUI. Dans un même ordre d'idée, l'étude de la phylogénie des bivalves et des mollusques en général apparaît comme une étape incontournable à l'avancement de l'étude de la DUI. Les connaissances déjà acquises sur ce système gagneraient à être positionnées dans un contexte phylogénétique plus rigoureux. Les principales perspectives de recherche qui apparaissent à l'issue de cette thèse sont, entre autres :

Prévalence de la DUI chez les bivalves

La DUI est extrêmement difficile à détecter, la confirmation de sa présence ou de son absence chez différents groupes de bivalves permettrait d'avoir une meilleure idée de sa distribution et permettrait d'en apprendre davantage sur son origine. Depuis la publication de la revue de littérature du chapitre II, la présence de la DUI a été confirmée chez un autre groupe de bivalve, les Donacidae (Theologidis *et al.*, 2008) et confirme la présence de la DUI dans quatre superfamilles de bivalves.

Séquençage de génomes complets mâle et femelle d'espèces dulcicoles et séquençage d'un génome nucléaire

L'effort de séquençage de génomes mitochondriaux mâle et femelle complets d'espèces d'Unionoïdés entrepris dans le chapitre III aurait tout avantage à être poursuivi. La comparaison des séquences de plusieurs génomes mitochondriaux complets pour un plus grand nombre d'espèces possédant la DUI pourrait permettre d'analyser la distribution de la variabilité génétique dans les génomes F et M, mais également d'analyser les changements structuraux survenus au cours de l'évolution du système. L'ordre des Unionoida inclut six familles : les Unionidae; les Margaritiferidae; les Hyriidae; les Etheriidae; les Mycetopodidae et les Iridinidea, le séquençage de génomes mitochondriaux complets provenant d'espèces de moules d'eau douce plus éloignées de celles séquencées dans cette thèse serait à privilégier. Ceci pourrait, par exemple, permettre de déterminer à quel moment dans l'évolution du système la Mcox2 extension a fait son apparition, ou encore, quelle est la lignée de bivalve dans laquelle la DUI est apparue en premier. Dans le même ordre d'idée, le séquençage de génomes complets d'espèces hermaphrodites permettrait possiblement de démontrer un lien entre le système de la DUI et le sexe.

Également, quoique plus complexe, le pyroséquençage de génome nucléaire complet d'une espèce d'Unionidés pour laquelle les génomes mitochondriaux mâle et femelle complets sont disponibles apparaît comme une étape charnière et incontournable. Une analyse comparative des transcrits selon différents tissus pourrait apporter une information permettant de mieux comprendre le rôle du génome M chez les unionidés et sa possible implication dans la détermination du sexe.

Les résultats de l'étude de génétique des populations du chapitre IV a permis de démontrer que différentes combinaisons d'haplotypes mitochondriaux M et F semblent pouvoir coexister. Ce résultat est une avenue particulièrement intéressante à explorer du point de vue de la coévolution cytonucléaire. Une comparaison plus fine des sections de gènes communes et la comparaison des différentes combinaisons pourraient permettre d'accumuler des indices de la quantité de divergence tolérée par le système et du même souffle sur la coévolution.

Les résultats du chapitre V offrent de nombreuses perspectives de recherche supplémentaires. Il serait possible entre autres, de tester en partie l'importance de la coévolution et des coadaptations en vérifiant si les sites de sélections positives sont corrélés chez les génomes mitochondriaux F et M. Si les génomes ont coévolués, on pourrait s'attendre à ce que les sites soient les mêmes ou dans les mêmes régions.

Les résultats obtenus dans le cadre de ce projet de doctorat ont permis de mettre en évidence le rôle de la sélection naturelle sur l'évolution moléculaire des génomes mitochondriaux mâles et femelles du système de la double transmission uniparentale chez les Unionoïdés. Toutefois, cette thèse

a aussi soulevé, des questions fondamentales. La première concerne la néofonctionnalisation probable de l'extension de la cytochrome c oxydase II. Cette extension est-elle, tel que précédemment suggéré, une protéine de la reproduction ? Les protéines de la reproduction sont de plus en plus étudiées chez différents organismes, mais les forces évolutives qui agissent sur celles-ci sont encore mal connues et jamais auparavant une telle protéine codée par le génome mitochondrial n'a été décrite.

Deuxièmement, une des questions soulevées par les résultats du chapitre V est pourquoi les forces sélectives semblent-elles agir différemment chez les bivalves unionidés que chez les autres groupes de bivalves possédant la DUI ? Les chapitres II et III fournissent une partie de la réponse et bien que certaines différences structurales existent entre les génomes M et F des différents groupes de bivalves, tout tend à démontrer que la présence de la Mcox2ext serait un facteur déterminant. La présence de cette extension et l'intense sélection positive observée sur l'ensemble des gènes du génome M pourraient être le reflet d'une nouvelle fonction complémentaire pour le génome M. Pour le moment, le rôle exact de cette extension fonctionnelle demeure inconnu, mais elle semble associée à la reproduction. D'un point de vue évolutif, la même question se pose : pourquoi l'apparition d'une telle extension chez les bivalves unionidés et chez aucun autre animal et ce, que leur système de transmission mitochondriale soit la SMI ou la DUI. La cox2extension est une nouveauté évolutive. Dans la discussion du chapitre III, il est proposé que les caractéristiques mitogénomiques observées chez les bivalves unionidés s'approchent davantage des conditions ancestrales de la DUI. Toutefois, à la lumière de l'ensemble des résultats de cette thèse, il m'apparaît plus parcimonieux de considérer que la cox2extension soit apparue seulement chez les unionidés plutôt que

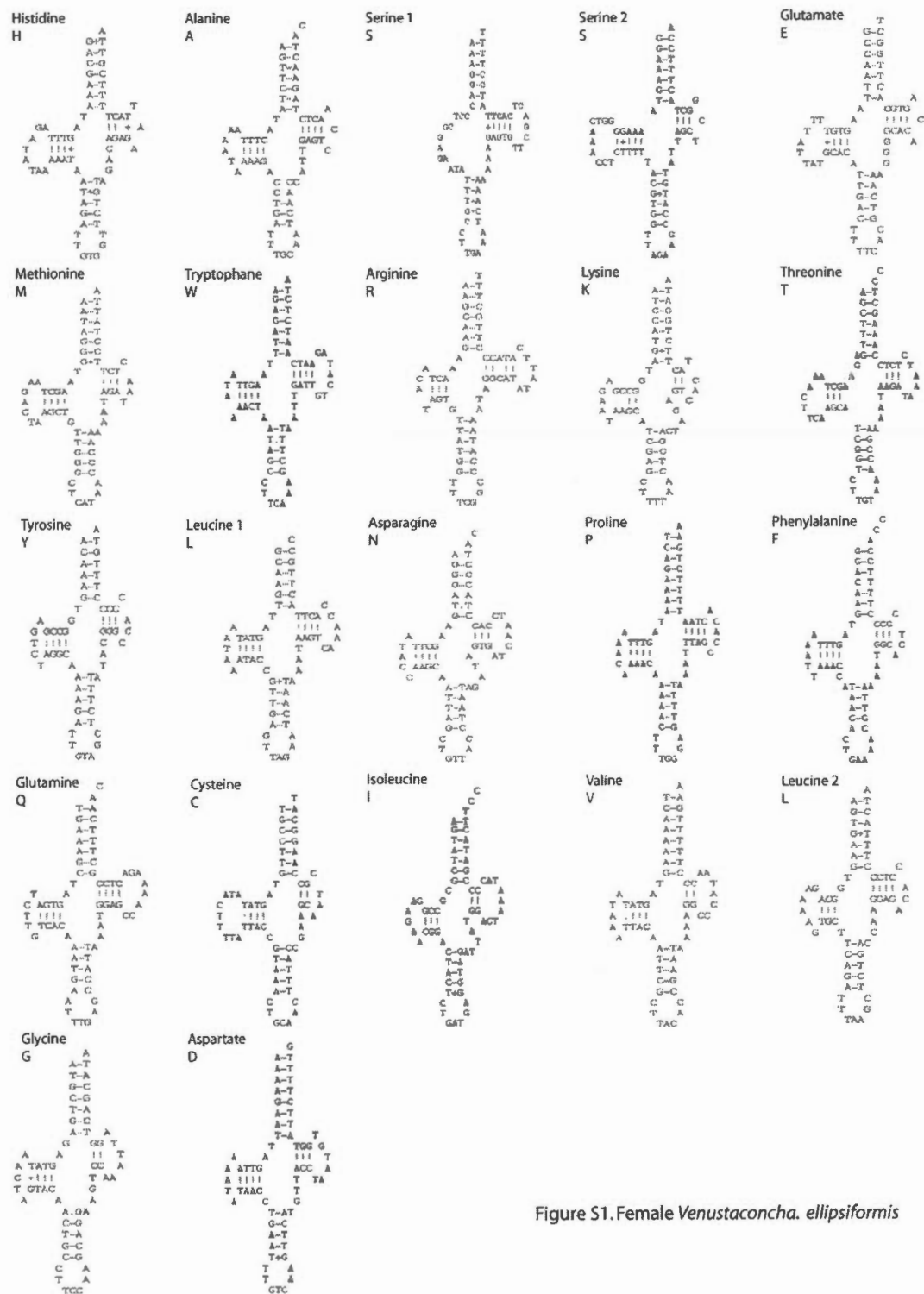
cette extension soit apparue chez un ancêtre des autolamellibranches et perdue par la suite chez les autres groupes de bivalves possédant la DUI. Le système de la DUI en modifiant les forces évolutives agissant sur le génome mitochondrial a fort probablement permis une relaxation des contraintes sur le génome M permettant l'apparition d'une nouveauté évolutive ce qui n'aurait pas été possible dans un système de transmission maternelle stricte. Nous sommes encore loin de pouvoir comprendre comment et pourquoi les bivalves Unionidés ont pu tirer avantage de cette nouveauté. Mais si la *cox2* extension est telle que proposée dans cette thèse, une nouveauté évolutive positivement sélectionnée issue d'une néofonctionnalisation suite à une duplication d'un génome mitochondrial, il m'apparaît prioritaire de poursuivre la recherche pour tenter d'en comprendre les fonctions. Pour trouver réponse à cette question, il m'apparaît impératif de revenir à l'étude de la biologie et de l'écologie des espèces concernées. Plusieurs facteurs associés à la reproduction des moules d'eau douce seraient à explorer (par exemple, polyandrie, compétition et collaboration spermatique) pour expliquer cette situation.

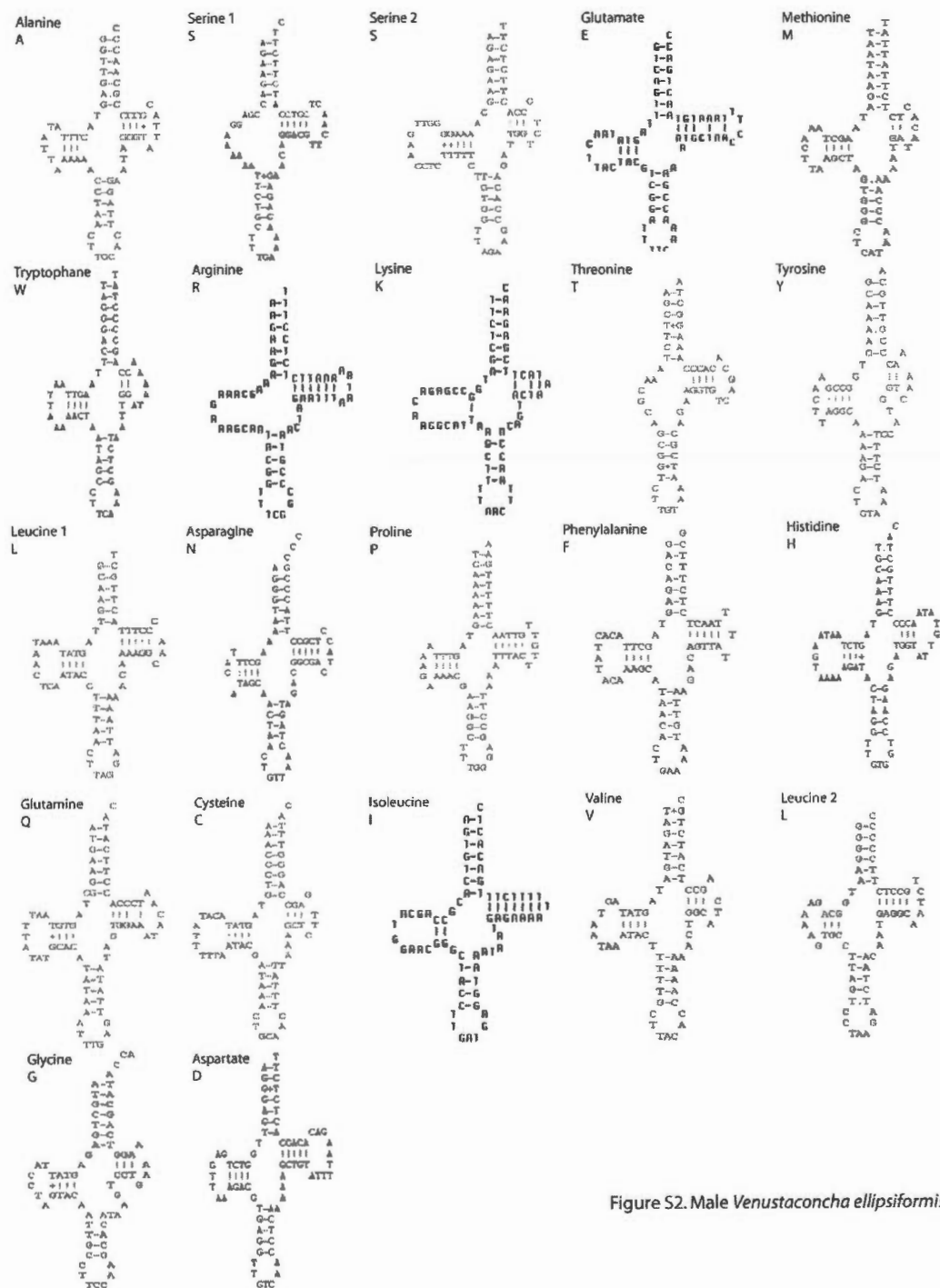
APPENDICE A

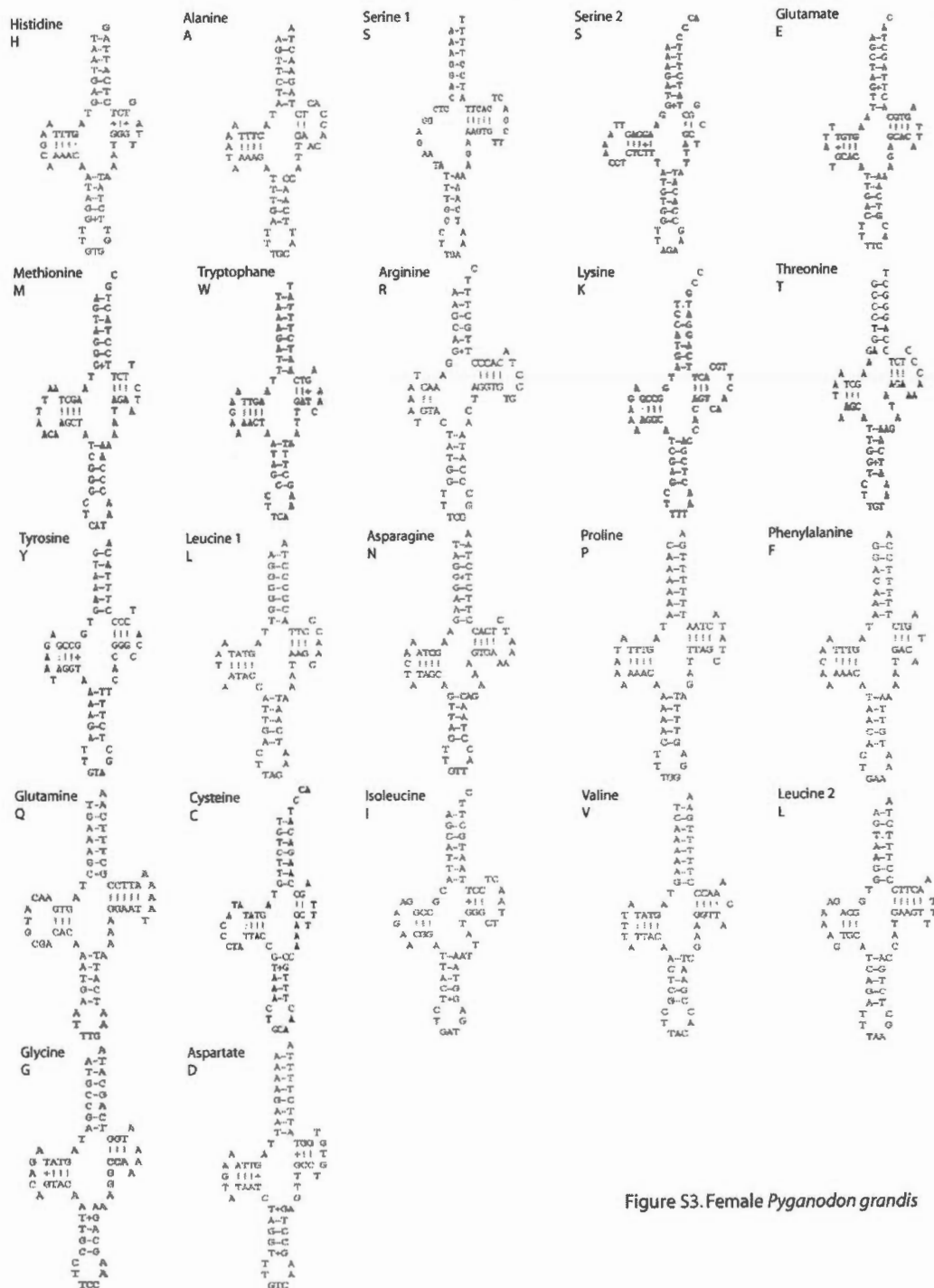
FIGURES SUPPLÉMENTAIRES DU CHAPITRE III

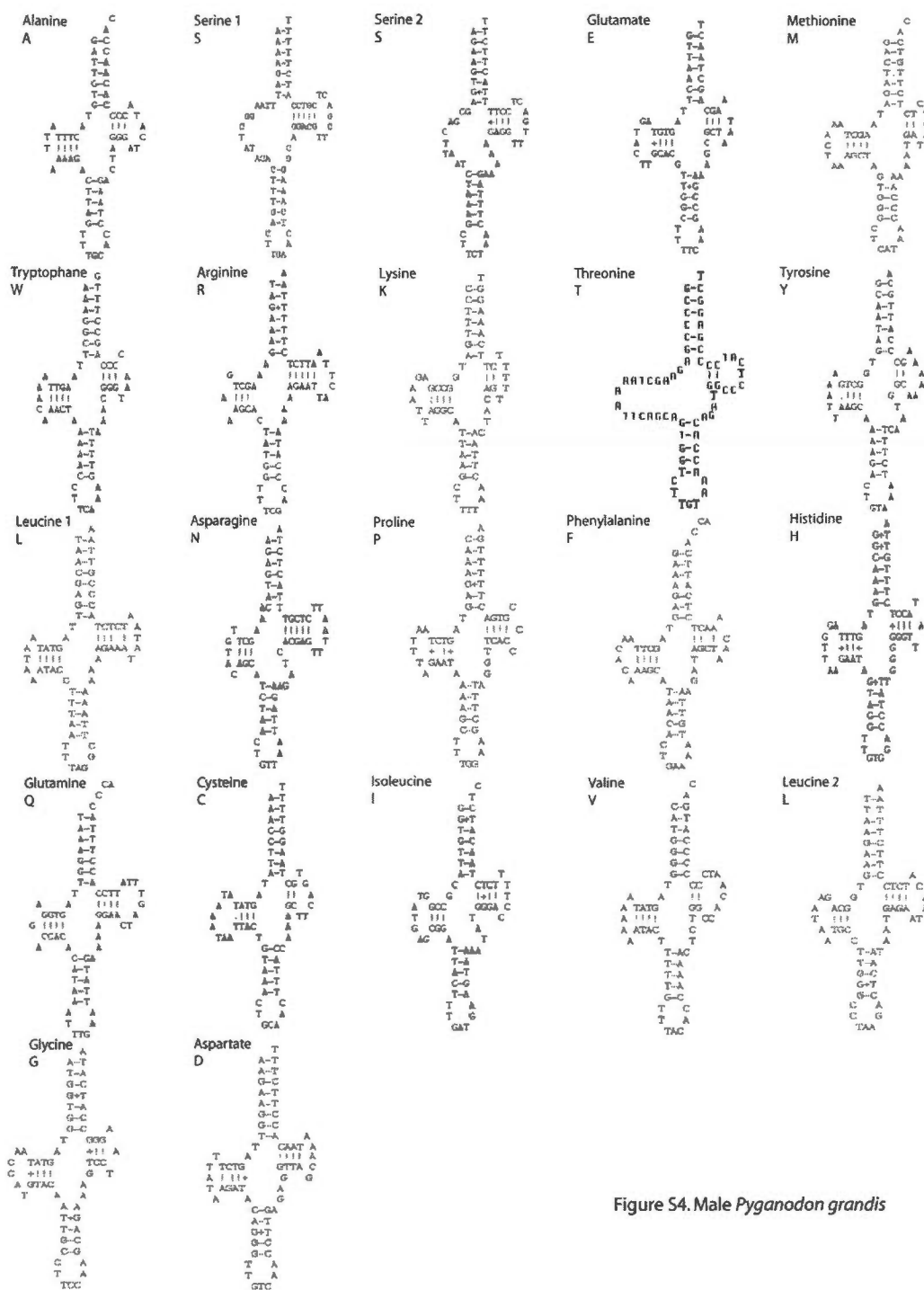
A.1	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>F Venustaconcha ellipsiformis</i> , shown in the order they occur in the genome, beginning with trnH. Amino acid identities are given above each sequence.....	220
A.2	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>M Venustaconcha ellipsiformis</i> , shown in the order they occur in the genome, beginning with trnA. Amino acid identities are given above each sequence.....	221
A.3	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>F Pyganodon grandis</i> , shown in the order they occur in the genome, beginning with trnH. Amino acid identities are given above each sequence.....	222
A.4	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>M Pyganodon grandis</i> , shown in the order they occur in the genome, beginning with trnA. Amino acid identities are given above each sequence.....	223
A.5	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>F Inversidens japonensis</i> , shown in the order they occur in the genome, beginning with trnH. Amino acid identities are given above each sequence.....	224
A.6	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>M Inversidens japonensis</i> , shown in the order they occur in the genome, beginning with trnA. Amino acid identities are given above each sequence.....	225
A.7	Inferred secondary structures of the 22 mitochondrial tRNAs from <i>F Quadrula quadrula</i> , shown in the order they occur in the genome, beginning with trnH. Amino acid identities are given	

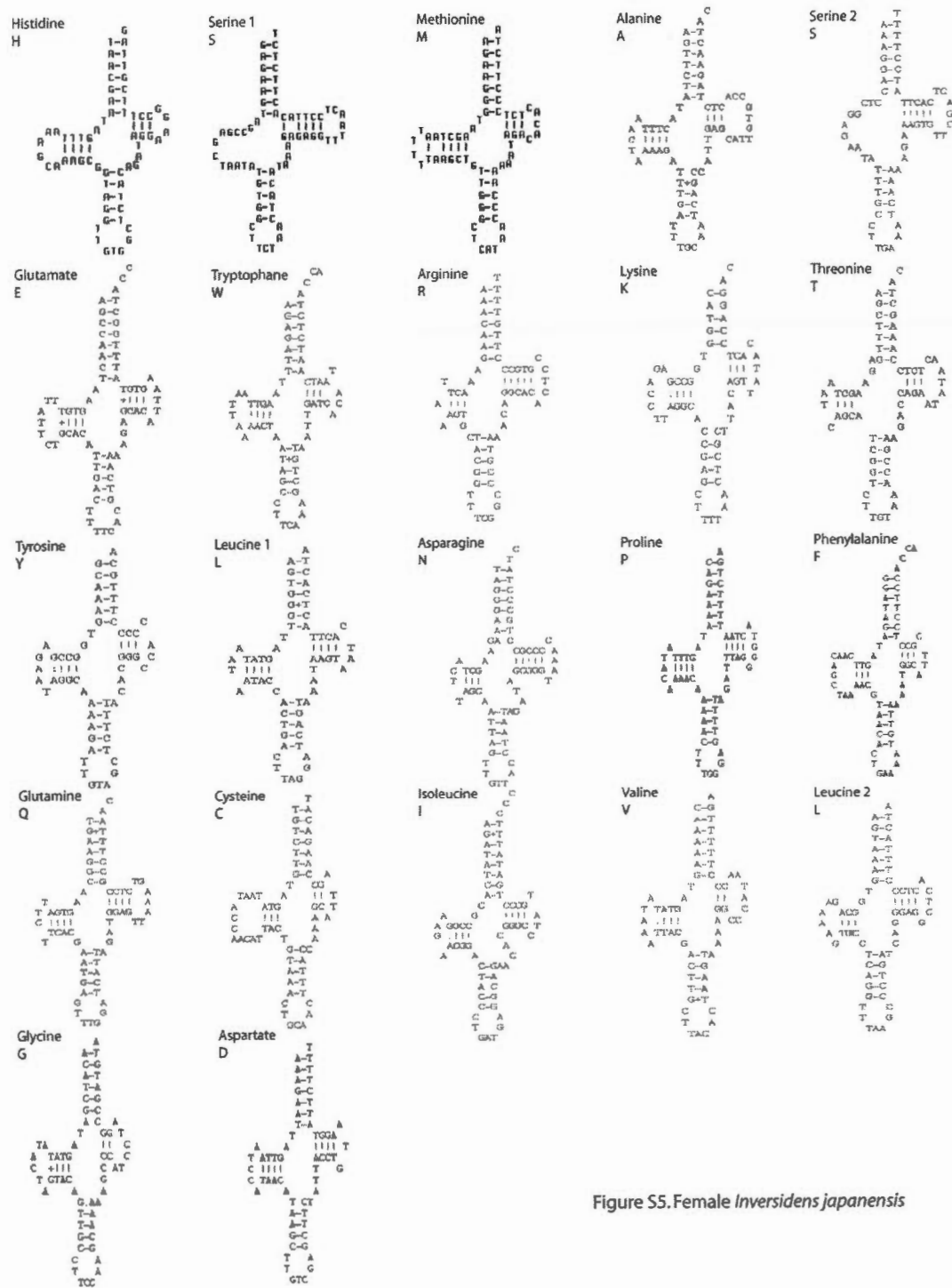
above each sequence.....	226
A.8 Inferred secondary structures of the 22 mitochondrial tRNAs from M <i>Quadrula quadrula</i> , shown in the order they occur in the genome, beginning with trnA. Amino acid identities are given above each sequence.....	227

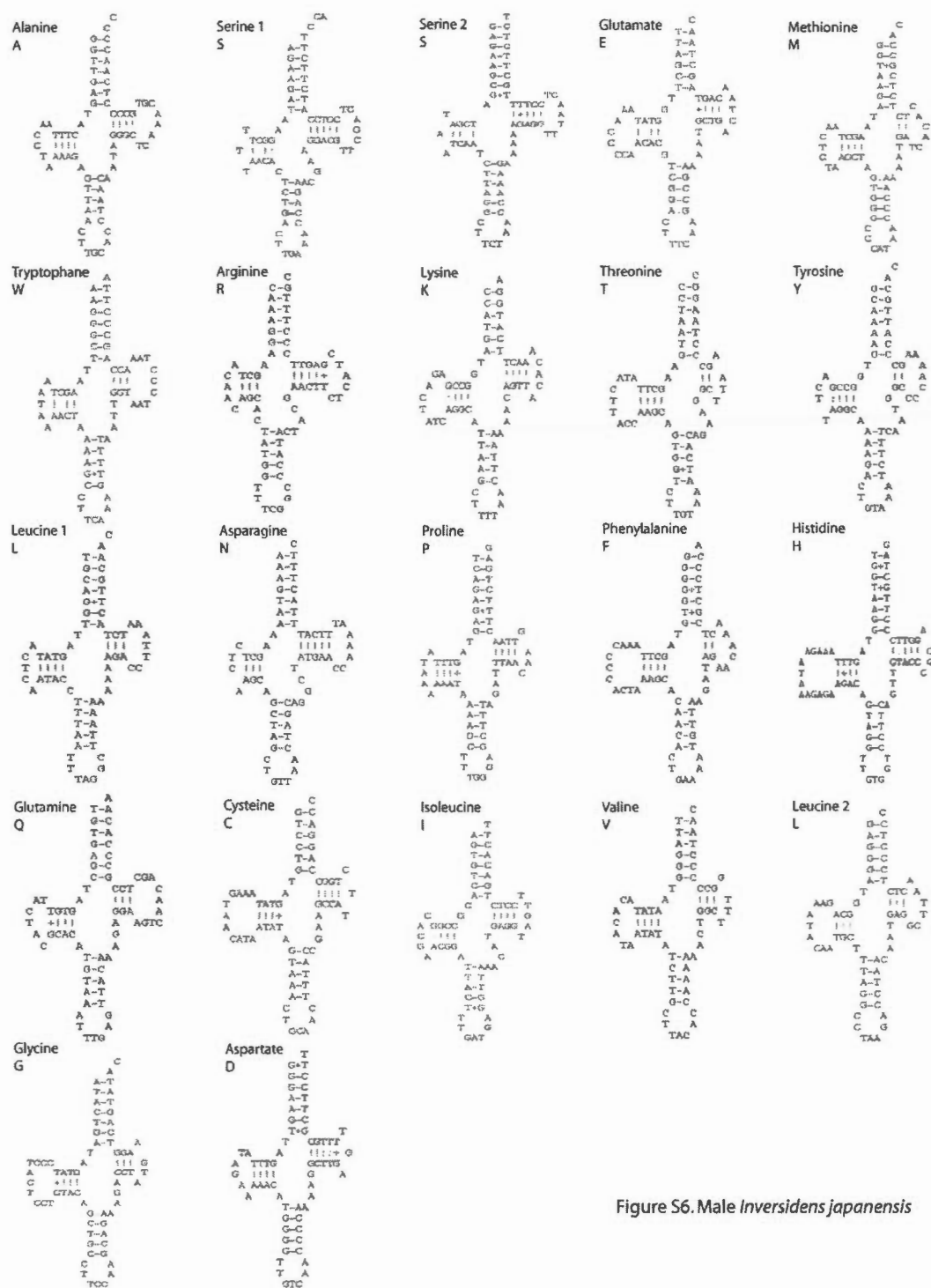
Figure S1. Female *Venustaconcha. ellipsiformis*

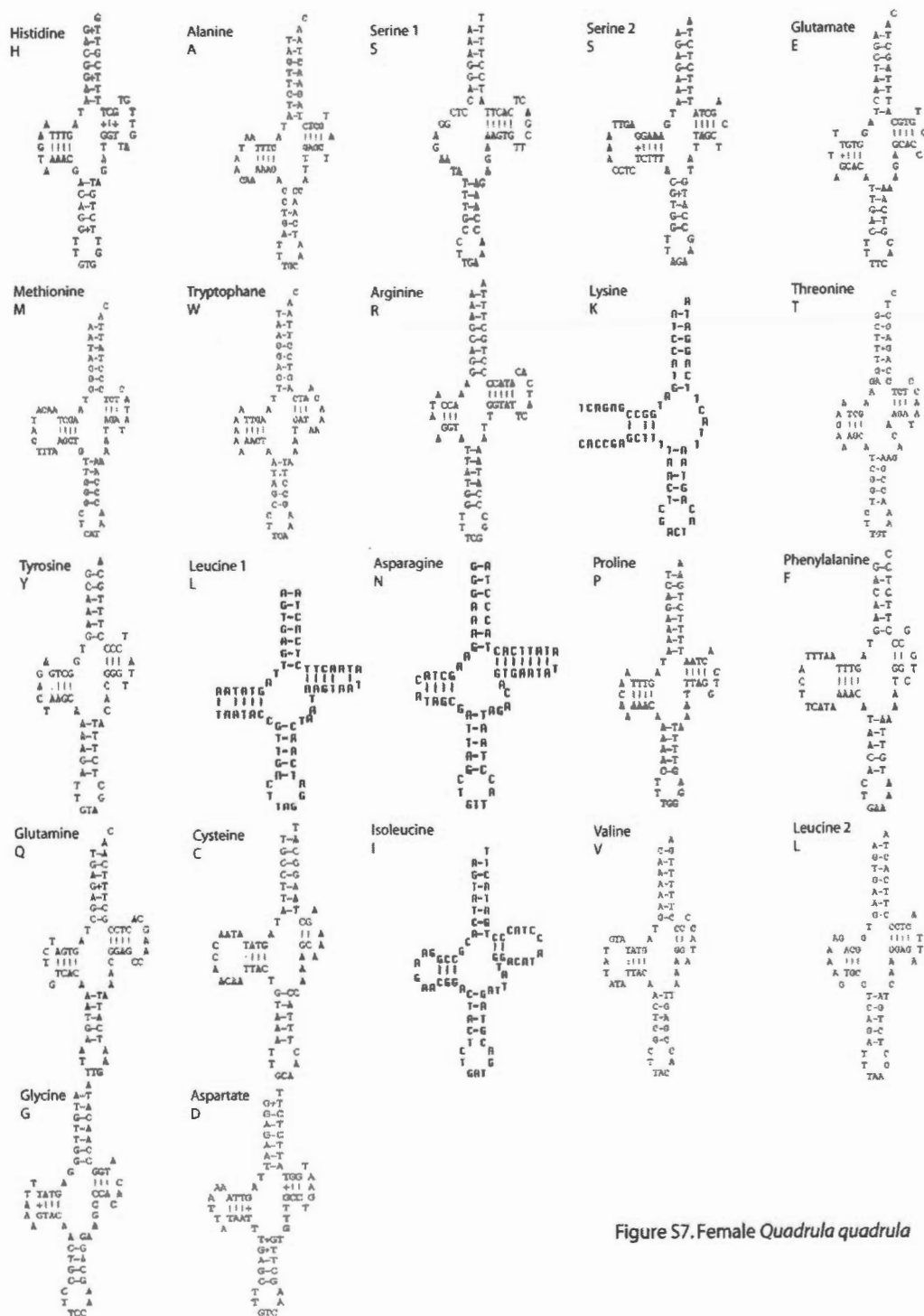
Figure S2. Male *Venustaconcha ellipsiformis*

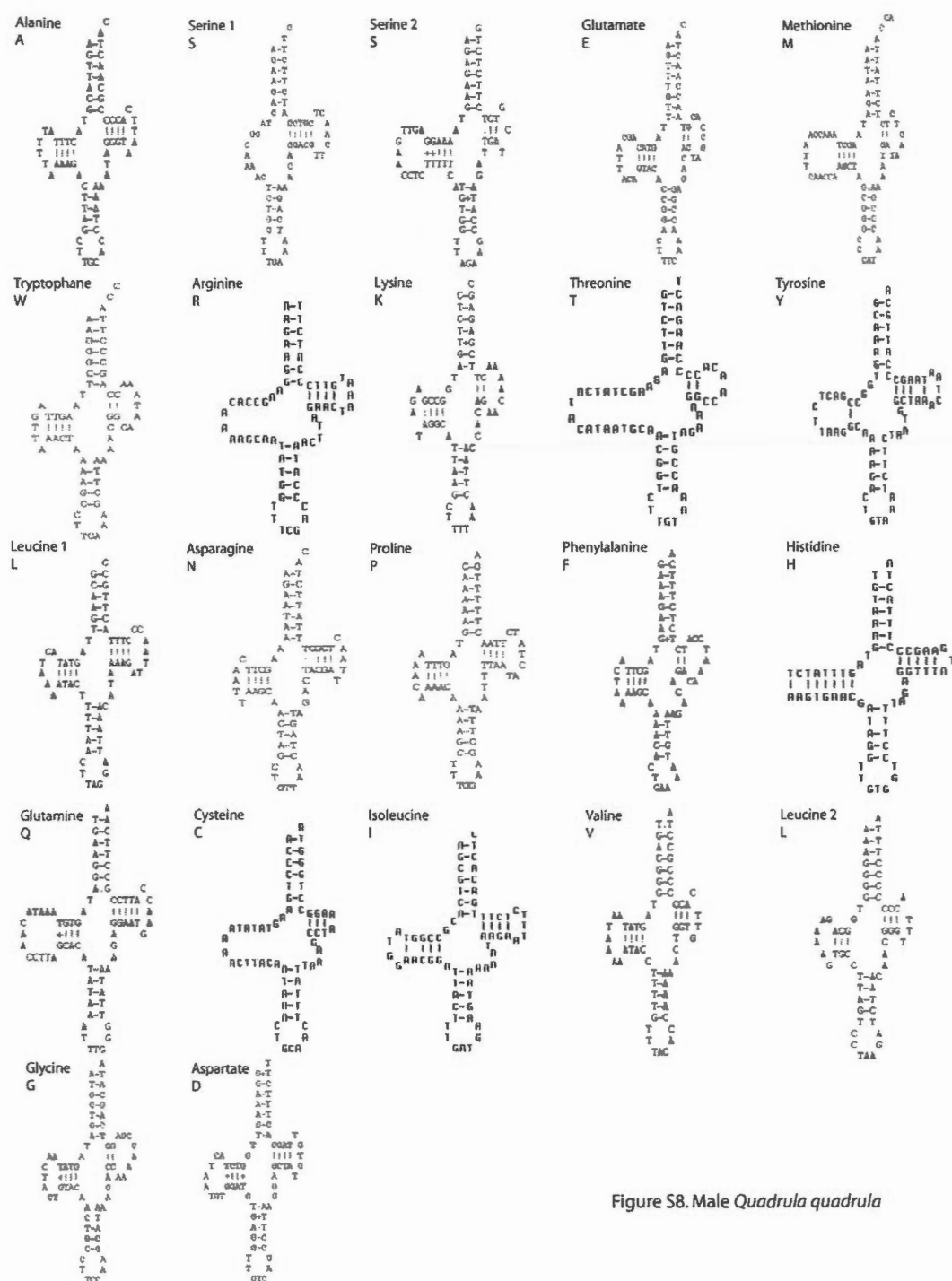
Figure S3. Female *Pyganodon grandis*

Figure S4. Male *Pyganodon grandis*

Figure S5. Female *Inversidens japonensis*

Figure S6. Male *Inversidens japonensis*



Figure S8. Male *Quadula quadrula*

APPENDICE B

TABLEAU SUPPLÉMENTAIRE DU CHAPITRE III

B.1	Codon usage in the female- and male-transmitted mitochondrial genomes of <i>Venustaconcha ellipsiformis</i> , <i>Pyganodon grandis</i> , <i>Inversidens japonensis</i> and <i>Quadrula quadrula</i>	229
-----	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Codon usage in the female- and male-transmitted mitochondrial genomes of *V. ellipsiformis*, *P. grandis*, *I. japonensis* and *Q. quadrula*.

	<i>V. ellipsiformis</i>				<i>V. ellipsiformis</i>				<i>P. grandis</i>				<i>I. japonensis</i>				<i>I. japonensis</i>				<i>Q. quadrula</i>			
	Female		Male		Female		Male		Female		Male		Female		Male		Female		Male		Female		Male	
AA	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Phe (F)	278	7.46	274	7.06	255	6.85	233	6.32	196	5.00	259	6.83	283	7.65	262	6.71								
	33	0.89	47	1.21	45	1.21	59	1.60	61	1.56	57	1.50	47	1.27	42	1.08								
Leu (L1)	187	5.02	157	4.05	231	6.20	137	3.71	86	2.19	217	5.72	168	4.54	157	4.02								
	184	4.94	195	5.02	128	3.44	188	5.10	144	3.68	122	3.22	205	5.54	191	4.89								
Ser (S2)	117	3.14	128	3.30	121	3.25	100	2.71	95	2.42	119	3.14	118	3.19	104	2.67								
	27	0.72	18	0.46	22	0.59	40	1.08	37	0.94	20	0.53	17	0.46	21	0.54								
	45	1.21	37	0.95	43	1.15	31	0.84	38	0.97	37	0.98	46	1.24	47	1.20								
	12	0.32	24	0.62	13	0.35	26	0.70	23	0.59	17	0.45	17	0.46	15	0.38								
Tyr (Y)	90	2.42	106	2.73	79	2.12	81	2.20	86	2.19	90	2.37	91	2.46	102	2.61								
	39	1.05	52	1.34	51	1.37	42	1.14	60	1.53	48	1.27	37	1.00	41	1.05								
TERM	5	0.13	9	0.23	8	0.21	7	0.19	75	1.91	8	0.21	9	0.24	6	0.15								
	7	0.19	3	0.08	5	0.13	5	0.14	52	1.33	2	0.05	4	0.11	4	0.10								
Cys (C)	50	1.34	85	2.19	41	1.10	41	1.11	69	1.76	69	1.82	50	1.35	87	2.23								
	11	0.30	8	0.21	14	0.38	18	0.49	30	0.77	25	0.66	10	0.27	14	0.36								
Trp (W)	39	1.05	43	1.11	57	1.53	38	1.03	55	1.40	68	1.79	44	1.19	52	1.33								
	75	2.01	69	1.78	58	1.56	68	1.84	108	2.76	43	1.13	71	1.92	64	1.64								
Leu (L2)	77	2.07	78	2.01	76	2.04	70	1.90	103	2.63	99	2.61	76	2.05	99	2.54								
	26	0.70	17	0.44	24	0.64	45	1.22	48	1.23	19	0.50	23	0.62	31	0.79								
	72	1.93	83	2.14	88	2.36	71	1.92	80	2.04	83	2.19	69	1.87	91	2.33								
	20	0.54	19	0.49	25	0.67	50	1.36	49	1.25	29	0.76	16	0.43	18	0.46								
Pro (P)	56	1.50	54	1.39	66	1.77	49	1.33	59	1.51	65	1.71	55	1.49	51	1.31								
	18	0.48	9	0.23	17	0.46	29	0.79	78	1.99	14	0.37	4	0.11	6	0.15								
	48	1.29	57	1.47	38	1.02	42	1.14	63	1.61	40	1.05	61	1.65	53	1.36								

	CGG	18	0.48	9	0.23	13	0.35	26	0.70	31	0.79	8	0.21	18	0.49	14	0.36
His (H)	CAT	62	1.66	49	1.26	56	1.50	45	1.22	49	1.25	58	1.53	54	1.46	60	1.54
	CAC	16	0.43	31	0.80	20	0.54	31	0.84	61	1.56	19	0.50	19	0.51	25	0.64
Gln (Q)	CAA	38	1.02	40	1.03	41	1.10	29	0.79	93	2.37	44	1.16	39	1.05	43	1.10
	CAG	24	0.64	22	0.57	20	0.54	32	0.87	38	0.97	23	0.61	23	0.62	24	0.62
Arg (R)	CGT	12	0.32	30	0.77	13	0.35	18	0.49	34	0.87	21	0.55	18	0.49	27	0.69
	CGC	6	0.16	3	0.08	5	0.13	7	0.19	10	0.26	2	0.05	3	0.08	6	0.15
	CGA	25	0.67	16	0.41	37	0.99	24	0.65	16	0.41	25	0.66	21	0.57	19	0.49
	CGG	16	0.43	12	0.31	6	0.16	16	0.43	21	0.54	9	0.24	21	0.57	10	0.26
Ile (I)	ATT	197	5.29	185	4.77	222	5.96	148	4.01	94	2.40	194	5.11	188	5.08	140	3.59
	ATC	49	1.32	53	1.37	52	1.40	56	1.52	58	1.48	49	1.29	54	1.46	70	1.79
Met (M)	ATA	108	2.90	117	3.01	118	3.17	79	2.14	74	1.89	140	3.69	91	2.46	117	3.00
	ATG	65	1.74	91	2.34	61	1.64	85	2.30	74	1.89	72	1.90	74	2.00	97	2.49
Thr (T)	ACT	73	1.96	86	2.22	74	1.99	71	1.92	65	1.66	83	2.19	61	1.65	81	2.08
	ACC	42	1.13	27	0.70	24	0.64	49	1.33	58	1.48	26	0.69	39	1.05	36	0.92
	ACA	50	1.34	65	1.67	62	1.66	39	1.06	62	1.58	59	1.56	53	1.43	65	1.67
	ACG	10	0.27	12	0.31	11	0.30	21	0.57	15	0.38	11	0.29	14	0.38	18	0.46
Asn (N)	AAT	57	1.53	67	1.73	60	1.61	52	1.41	65	1.66	70	1.85	54	1.46	54	1.38
	AAC	30	0.81	44	1.13	41	1.10	41	1.11	80	2.04	41	1.08	38	1.03	41	1.05
Lys (K)	AAA	50	1.34	48	1.24	58	1.56	43	1.17	75	1.91	64	1.69	42	1.14	47	1.20
	AAG	42	1.13	49	1.26	32	0.86	35	0.95	48	1.23	36	0.95	39	1.05	48	1.23
Ser (S1)	AGT	56	1.50	71	1.83	61	1.64	35	0.95	42	1.07	54	1.42	60	1.62	74	1.90
	AGC	12	0.32	12	0.31	20	0.54	23	0.62	49	1.25	31	0.82	15	0.41	21	0.54
	AGA	36	0.97	54	1.39	60	1.61	26	0.70	39	1.00	84	2.21	36	0.97	40	1.03
	AGG	51	1.37	59	1.52	49	1.32	66	1.79	43	1.10	66	1.74	49	1.32	62	1.59
Val (V)	GTT	190	5.10	224	5.77	148	3.97	170	4.61	136	3.47	139	3.66	186	5.03	226	5.79
	GTC	16	0.43	20	0.52	27	0.72	41	1.11	26	0.66	20	0.53	20	0.54	20	0.51
	GTA	73	1.96	83	2.14	96	2.58	64	1.73	58	1.48	121	3.19	62	1.68	88	2.26
	GTG	91	2.44	104	2.68	68	1.83	110	2.98	97	2.48	58	1.53	100	2.70	107	2.74

Ala (A)	GCT	106	2.85	84	2.16	113	3.03	111	3.01	63	1.61	85	2.24	131	3.54	89	2.28
	GCC	35	0.94	14	0.36	33	0.89	61	1.65	44	1.12	26	0.69	26	0.70	33	0.85
	GCA	47	1.26	41	1.06	52	1.40	38	1.03	44	1.12	62	1.63	42	1.14	49	1.26
	GCG	19	0.51	13	0.33	12	0.32	32	0.87	21	0.54	20	0.53	13	0.35	9	0.23
Asp (D)	GAT	58	1.56	70	1.80	49	1.32	52	1.41	43	1.10	52	1.37	58	1.57	64	1.64
	GAC	12	0.32	13	0.33	15	0.40	16	0.43	20	0.51	15	0.40	11	0.30	16	0.41
Glu (E)	GAA	30	0.81	35	0.90	51	1.37	23	0.62	41	1.05	57	1.50	38	1.03	30	0.77
	GAG	52	1.40	52	1.34	44	1.18	56	1.52	48	1.23	46	1.21	50	1.35	55	1.41
Gly (G)	GGT	119	3.19	122	3.14	101	2.71	89	2.41	88	2.25	68	1.79	116	3.14	117	3.00
	GGC	25	0.67	14	0.36	23	0.62	49	1.33	41	1.05	18	0.47	23	0.62	12	0.31
	GGA	51	1.37	60	1.55	68	1.83	47	1.27	33	0.84	92	2.43	62	1.68	54	1.38
	GGG	140	3.76	108	2.78	104	2.79	163	4.42	126	3.22	75	1.98	117	3.16	136	3.49

APPENDICE C

FIGURES SUPPLÉMENTAIRES DU CHAPITRE V

- C.1 Amino acid alignments at positively selected sites based on BEB inference over male Unionidae. The intensity of the column color reflects the respective probability for each site to be positively selected..... 233
- C.2 Amino acid alignments at positively selected sites based on BEB inference over female Unionidae. The intensity of the column color reflects the respective probability for each site to be positively selected..... 234
- C.3 Amino acid alignments at positively selected sites based on BEB inference over female Verenidae. The intensity of the column color reflects the respective probability for each site to be positively selected..... 235
- C.4 Amino acid alignments at positively selected sites based on BEB inference over male Verenidae. The intensity of the column color reflects the respective probability for each site to be positively selected..... 236

	COI	COII	COIII	1F	NO1	NO2	NO3	NO4	NO5	CO8	
NC 01518 <i>Ca. Nostoc</i>	A-VGLLVNSV.NSLTTALD	QWSTSF	IANIGTFSH-CL	F	Y EFP	MFYKTD	KPVL	VGYSK N-LLS	NS-LWPI-SALASASH	ME-LSBKNDP	Ostreoida
NC 01519 <i>Ca. Nostoc</i>	A-VGLLVNSV.NSLTTALD	QWSTSF	IANIGTFSH-CL	F	Y EIF	ELNFTKTD	KPVL	VAFSR S-LHR	NS-LWPI-SALASASH	ME-LSBKNDP	
NC 01520 <i>Ca. Nostoc</i>	A-VGLLVNSV.NSLTTAMD	QWSTSF	IANIGTFSH-CL	F	Y EGF	ELNFTKTD	KPVL	VGYSK S-LHR	NS-LWPI-SALASASH	ME-LSBKNDP	
NC 01518 <i>Ca. Nostoc</i>	W-GFLVP3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPA-	VLVMS M-CH	V5-GGVSAI-LCAML	ME-LVYDNQV	Pectinoida
NC 01519 <i>Ca. Nostoc</i>	W-GALVP3SGV.SAFVAA	-SYBAGL	AHEAKCHACL	E	Y ELF	GFR3T0	KPA-	LVMS S-CH	V5-GGVSAVVC-ILM	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VWV13 G-CH	V5-GGVSAVVC-ILM	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VWV13 G-CH	V5-GGVSAVVC-ILM	ME-LVYDNQV	Other Veneridae
NC 01519 <i>Ca. Nostoc</i>	S-APLV3SGV.SAFVAA	LYTSKLL	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Veneridae F
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Veneridae M
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Unionidae F
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Unionidae M
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Mytilidae F
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Mytilidae M
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01518 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	Outgroups
NC 01519 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	
NC 01520 <i>Ca. Nostoc</i>	W-GPLT3SGV.SAFVAA	QWTSML	AHDSKGHACL	W	Y ELF	GFR3T0	KPAL	VLVSA S-CH	L5-GGVSAVVC-ILL	ME-LVYDNQV	

**Amino acid alignments at positively selected sites based on BEB inference over Female Unionidae
(30 sites shown with posteriori probability above 0.95)**

	COI	COII	COIII	ATP6	ND1	ND2	ND3	ND4	ND4L	ND5	COB																						
NC_01518 Clu haplogroup 1	V	-	D	F	F	I	F	H	T	S	G	L	W	M	S	C	S	E	C	F	M	F	F	L	S	F	D	D	E	S	Ostreoida		
NC_01715 Clu haplogroup 2	V	-	D	F	F	I	F	H	T	S	G	L	V	L	S	C	M	E	C	F	M	F	F	L	S	F	D	D	Q	E	S	Ostreoida	
NC_01726 Clu haplogroup 3	V	-	D	F	F	I	F	H	T	S	G	L	W	M	T	C	S	E	C	F	M	F	F	L	S	F	D	D	Q	E	S	Ostreoida	
NC_01518 Clu haplogroup 4	K	K	C	V	I	I	W	A	S	G	S	F	V	N	R	G	C	L	V	F	C	L	F	F	S	I	L	D	T	T	S	Pectinoida	
NC_01713 Clu haplogroup 5	K	K	C	V	I	I	W	A	S	G	S	F	V	N	R	G	C	L	V	F	C	L	F	F	S	I	L	D	T	T	S	Pectinoida	
NC_01713 Clu haplogroup 6	K	K	C	V	I	I	W	A	S	G	S	F	V	N	R	G	C	L	V	F	C	L	F	F	S	I	L	D	T	T	S	Pectinoida	
NC_01713 Clu haplogroup 7	K	K	C	V	I	I	W	A	S	G	S	F	V	N	R	G	C	L	V	F	C	L	F	F	S	I	L	D	T	T	S	Pectinoida	
NC_01713 Clu haplogroup 8	L	K	-	-	-	-	V	W	A	S	S	F	V	K	R	S	C	E	V	V	S	C	L	L	S	I	M	D	T	S	S	Pectinoida	
NC_01518 Clu haplogroup 9	L	K	-	-	-	-	V	W	A	S	S	F	V	K	R	S	C	E	V	V	S	C	L	L	S	I	M	D	T	S	S	Pectinoida	
NC_01518 Clu haplogroup 10	L	K	-	-	-	-	V	W	A	S	S	F	V	K	R	S	C	E	V	V	S	C	L	L	S	I	M	D	T	S	S	Pectinoida	
NC_01518 Clu haplogroup 11	G	E	N	A	Y	I	L	T	T	F	S	F	L	E	K	L	A	K	L	L	A	S	L	L	M	A	S	E	S	S	G	Other Veneridae	
NC_01518 Clu haplogroup 12	S	W	S	S	Y	V	L	C	L	A	V	S	G	V	G	I	S	L	L	M	S	L	L	M	A	S	S	Q	S	S	G	Other Veneridae	
NC_01518 Clu haplogroup 13	K	E	N	Y	F	C	L	A	L	H	G	F	V	S	G	M	S	L	L	M	S	L	L	M	A	S	E	Q	S	S	T	S	Other Veneridae
NC_01518 Clu haplogroup 14	G	L	-	C	F	M	W	K	T	F	S	F	V	E	V	V	S	K	Y	M	L	G	L	L	A	S	E	D	S	G	Veneridae F		
NC_01518 Clu haplogroup 15	G	L	S	C	F	M	W	K	T	F	S	F	A	E	V	V	T	X	G	N	L	L	S	S	A	S	E	D	S	G	Veneridae M		
NC_01518 Clu haplogroup 16	A	G	L	S	C	F	M	W	K	T	F	S	F	A	E	V	V	T	X	G	N	L	L	S	S	A	S	E	D	S	G	Veneridae M	
NC_01518 Clu haplogroup 17	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 18	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 19	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 20	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 21	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 22	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 23	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 24	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 25	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 26	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 27	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 28	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 29	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 30	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 31	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 32	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 33	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 34	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 35	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 36	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 37	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 38	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 39	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 40	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 41	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 42	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 43	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 44	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 45	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 46	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 47	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 48	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 49	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 50	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 51	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 52	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 53	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 54	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 55	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 56	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 57	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 58	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 59	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 60	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 61	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 62	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 63	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 64	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V	F	F	Y	T	T	S	G	G	Y	S	S	S	Unionidae F	
NC_01518 Clu haplogroup 65	P	V	A	Y	M	L	T	H	H	Q	M	C	A	S	S	A	I	S	V														

BIBLIOGRAPHIE

- Alacs, E., A. Georges, N. FitzSimmons et J. Robertson. 2010. «DNA detective: A review of molecular approaches to wildlife forensics». *Forensic Science, Medicine, and Pathology*, vol. 6, no 3, p. 180-194.
- Albu, M., X. J. Min, D. Hickey et B. Golding. 2008. «Uncorrected nucleotide bias in mtDNA can mimic the effects of positive darwinian selection». *Molecular Biology and Evolution*, vol. 25, no 12, p. 2521-2524.
- Amer, S. A. M. et Y. Kumazawa. 2007. «The mitochondrial genome of the lizard *Calotes versicolor* and a novel gene inversion in south asian draconine agamids». *Molecular Biology and Evolution*, vol. 24, no 6, p. 1330-1339.
- Anisimova, M., J. P. Bielawski et Z. Yang. 2001. «Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution». *Molecular Biology and Evolution*, vol. 18, no 8, p. 1585-1592.
- Anisimova, M. et D. A. Liberles. 2007. «The quest for natural selection in the age of comparative genomics». *Heredity*, vol. 99, no 6, p. 567-579.
- Attardi, G. 1985. «Animal mitochondrial DNA: An extreme example of genetic economy.». *International Review of Cytology*, vol. 93, no, p. 93-145.
- Awise, J. C. 1994. *Molecular markers, natural history and evolution*. New York: Chapman and Hall.
- Ballard, J. W. O. et M. C. Whitlock. 2004. «The incomplete natural history of mitochondria». *Molecular Ecology*, vol. 13, no 4, p. 729-744.
- Barber, D. C., A. Dyke, C. Hillaire-Marcel, A. E. Jennings, J. T. Andrews, M. W. Kerwin, G. Bilodeau, R. McNeely, J. Southon, M. D. Morehead et J. M. Gagnon. 1999. «Forcing of the cold event of 8,200 years ago by catastrophic drainage of laurentide lakes». *Nature*, vol. 400, no 6742, p. 344-348.
- Barrientos, A., L. Kenyon et C. T. Moraes. 1998. «Human xenomitochondrial cybrids - cellular models of mitochondrial complex i deficiency». *The Journal of Biological Chemistry*, vol. 273, no 23, p. 14210-14217.

- Barrowclough, G. F. et R. M. Zink. 2009. «Funds enough, and time: mtDNA, nuDNA and the discovery of divergence». *Molecular Ecology*, vol. 18, no 14, p. 2934-2936.
- Bassham, S., C. Canestro et J. Postlethwait. 2008. «Evolution of developmental roles of pax2/5/8 paralogs after independent duplication in urochordate and vertebrate lineages». *BMC Biology*, vol. 6, no 35, doi: 10.1186/1741-7007-6-35.
- Bauer, G. 1987. «Reproductive strategy of the freshwater pearl mussel *Margaritifera margaritifera*». *Journal of Animal Ecology*, vol. 56, no 2, p. 691-704.
- Bauer, G. et K. Wächtler. 2000. *Ecology and evolution of the freshwater mussels unionoida*. Berlin: Springer-Verlag.
- Bazin, E., S. Glemin et N. Galtier. 2006. «Population size does not influence mitochondrial genetic diversity in animals». *Science*, vol. 312, no 5773, p. 570-572.
- Beagley, C. T., K. A. Taylor et D. R. Wolstenholme. 1997. «Gender-associated diverse mitochondrial DNA molecules of the mussel *Mytilus californianus*». *Current Genetics*, vol. 31, no 4, p. 318-324.
- Bendall, K. E., V. A. Macaulay et B. C. Sykes. 1997. «Variable levels of a heteroplasmic point mutation in individual hair roots». *American Journal of Human Genetics*, vol. 61, no 6, p. 1303-1308.
- Bieler, R., J.G. Carter, et E.V. Coan. 2010. «Classification of bivalve families» In *Nomenclator of Bivalve Families with a Classification of Bivalve Families*, P. Bouchet et J.-P. Rocroi (éd.), Malacologia, vol. 52, no 2, p. 113-133.
- Bilton, D. T., J. R. Freeland et B. Okamura. 2001. «Dispersal in freshwater invertebrates». *Annual Review of Ecology and Systematics*, vol. 32, p. 159-181.
- Birky, C. W. 1995. «Uniparental inheritance of mitochondrial and chloroplast genes - mechanisms and evolution». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no 25, p. 11331-11338.

- Birky, C. W. 2001. «The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models». *Annual Review of Genetics*, vol. 35, p. 125.
- Birky, C. W., T. Maruyama et P. Fuerst. 1983. «An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results». *Genetics*, vol. 103, no 3, p. 513-527.
- Blier, P. U., et H. Lemieux. 2001. «The impact of the thermal sensitivity of cytochrome c oxidase on the respiration rate of arctic charr red muscle mitochondria». *Journal of Comparative Physiology B- Biochemical, Systemic and Environmental Physiology*, vol. 171, no 3, p. 247-253.
- Blier, P. U., S. Breton, V. Desrosiers et H. Lemieux. 2006. «Functional conservatism in mitochondrial evolution: Insight from hybridization of arctic and brook charrs». *Journal of Experimental Zoology Part B- Molecular and Developmental Evolution*, vol. 306B, no 5, p. 425-432.
- Blier, P. U., F. Dufresne et R. S. Burton. 2001. «Natural selection and the evolution of mtDNA-encoded peptides: Evidence for intergenomic co-adaptation». *Trends in Genetics*, vol. 17, no 7, p. 400-406.
- Bogan, A.E., 1993. «Freshwater bivalve extinctions (Mollusca: Unionoida): a search for causes». *American Zoologist*, vol 33, p. 599-60.
- Bogan, A. E. et W. R. Hoeh. 2000. «On becoming cemented: Evolutionary relationships among the genera in the freshwater bivalve family Etheriidae (Bivalvia: Unionoida)». *Geological Society, London, Special Publications*, vol. 177, no 1, p. 159-168.
- Bogenhagen, D. F. 1999. «Repair of mtDNA in vertebrates». *The American Journal of Human Genetics*, vol. 64, no 5, p. 1276-1281.
- Bogenhagen, D. F. et D. A. Clayton. 2003. «The mitochondrial DNA replication bubble has not burst». *Trends in Biochemical Sciences*, vol. 28, no 7, p. 357-360.
- Boore, J. L. 1999. «Animal mitochondrial genomes». *Nucleic Acids Research*, vol. 27, no 8, p. 1767-1780.

- Boore, J. L. 2006. «The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda)». *BMC Genomics*, vol. 7, no 182, doi: 10.1186/1471-2164-7-182.
- Boore, J. L., L. L. Daehler et W. M. Brown. 1999. «Complete sequence, gene arrangement, and genetic code of mitochondrial DNA of the cephalochordate *Branchiostoma floridae* (amphioxus)». *Molecular Biology and Evolution*, vol. 16, no 3, p. 410-418.
- Boore, J. L., D. V. Lavrov et W. M. Brown. 1998. «Gene translocation links insects and crustaceans». *Nature*, vol. 392, no 6677, p. 667.
- Boore, J. L., M. Medina et L. A. Rosenberg. 2004. «Complete sequences of the highly rearranged molluscan mitochondrial genomes of the scaphopod *Graptacme eborea* and the bivalve *Mytilus edulis*». *Molecular Biology and Evolution*, vol. 21, no 8, p. 1492-1503.
- Bowmaker, M., M. Y. Yang, T. Yasukawa, A. Reyes, H. T. Jacobs, J. A. Huberman et I. J. Holt. 2003. «Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone». *Journal of Biological Chemistry*, vol. 278, no 51, p. 50961-50969.
- Brake, J., J. Davidson et J. Davis. 2004. «Field observations on growth, gametogenesis, and sex ratio of triploid and diploid *Mytilus edulis*». *Aquaculture*, vol. 236, no 1-4, p. 179-191.
- Breton, S. et H. Doucet-Beaupré. 2007. «A sexually fair system of mitochondrial DNA transmission». *M S-Medecine Sciences*, vol. 23, no 11, p. 1038-1040.
- Breton, S., H. Doucet-Beaupré, D. T. Stewart, W. R. Hoeh and P. U. Blier. 2007. «The unusual system of doubly uniparental inheritance of mtDNA: Isn't one enough?». *Trends In Genetics*, vol. 23, no 9, p. 465-474.
- Breton, S., H. Doucet-Beaupré, D. T. Stewart, H. Piontkivska, M. Karmakar, A. E. Bogan, P. U. Blier et W. R. Hoeh. 2009. «Comparative mitochondrial genomics of freshwater mussels (Bivalvia: Unionoida) with doubly uniparental inheritance of mtDNA: Gender-specific open reading frames and putative origins of replication». *Genetics*, vol. 183, no 4, p. 1575-1589.

- Breton, S., G. Burger, D. T. Stewart et P. U. Blier. 2006. «Comparative analysis of gender-associated complete mitochondrial genomes in marine mussels (*Mytilus* spp.)». *Genetics*, vol. 172, no 2, p. 1107-1119.
- Breton, S., D. T. Stewart et P. U. Blier. 2009. «Role-reversal of gender-associated mitochondrial DNA affects mitochondrial function in *Mytilus edulis* (Bivalvia: Mytilidae)». *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution*, vol. 312B, no 2, p. 108-117.
- Breton, S., D. T. Stewart, S. Shepardson, R. J. Trdan, A. E. Bogan, E. G. Chapman, A. J. Ruminas, H. Piontkivska et W. R. Hoeh. 2011. «Novel protein genes in animal mtDNA: A new sex determination system in freshwater mussels (Bivalvia: Unionoida)?». *Molecular Biology and Evolution*, vol. 28, p. 1645-1659.
- Brower, A. V. Z. 1999. «Delimitation of phylogenetic species with DNA sequences: A critique of Davis and Nixon's population aggregation analysis». *Systematic Biology*, vol. 48, no 1, p. 199-213.
- Buhay, J. E., J. M. Serb, C. R. Dean, Q. Parham et C. Lydeard. 2002. «Conservation genetics of two endangered unionid bivalve species, *Epioblasma florentina walkeri* and *E. capsaeformis* (Unionidae: Lampsilini)». *Journal of Molluscan Studies*, vol. 68, no 4, p. 385-391.
- Burch, J. B. 1975. *Freshwater unionacean clams (Mollusca: Pelecypoda) of North America*. Hamburg (Michigan): University of Michigan Library.
- Burger, G., M. W. Gray et B. F. Lang. 2003. «Mitochondrial genomes: Anything goes». *Trends in Genetics*, vol. 19, no 12, p. 709-716.
- Burgstaller, J., P. Schinogl, A. Dinnyes, M. Muller et R. Steinborn. 2007. «Mitochondrial DNA heteroplasmy in ovine fetuses and sheep cloned by somatic cell nuclear transfer». *BMC Developmental Biology*, vol. 7, no 141, doi : 10.1186/1471-213X-7-141.
- Burky, A. J. 1983. «Physiological ecology of freshwater bivalves». In *The Mollusca*, W. D. Russell-Hunter (éd.). Orlando, Florida: Academic Press.
- Burt, A. et R. L. Trivers. 2006. *Genes in conflict*. Boston (Massachusetts): Belknap Press of Harvard University Press.

- Burton, R. S., P. D. Rawson et S. Edmands. 1999. «Genetic architecture of physiological phenotypes: Empirical evidence for coadapted gene complexes». *American Zoologist*, vol. 39, no 2, p. 451-462.
- Burzynski, A., M. Zbawicka, D. O. F. Skibinski et R. Wenne. 2003. «Evidence for recombination of mtDNA in the marine mussel *Mytilus trossulus* from the baltic». *Molecular Biology and Evolution*, vol. 20, no 3, p. 388-392.
- Burzynski, A., M. Zbawicka, D. O. F. Skibinski et R. Wenne. 2006. «Doubly uniparental inheritance is associated with high polymorphism for rearranged and recombinant control region haplotypes in baltic *Mytilus trossulus*». *Genetics*, vol. 174, no 3, p. 1081-1094.
- Cai, J. J. 2008. «Pgetoolbox: A matlab toolbox for population genetics and evolution». *Journal of Heredity*, vol. 99, no 4, p. 438-440.
- Cantatore, P., M. N. Gadaleta, M. Roberti, C. Saccone et A. C. Wilson. 1987. «Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes». *Nature*, vol. 329, no 6142, p. 853.
- Cao, L. Q., E. Kenchington et E. Zouros. 2004. «Differential segregation patterns of sperm mitochondria in embryos of the blue mussel (*Mytilus edulis*)». *Genetics*, vol. 166, no 2, p. 883-894.
- Cao, L. Q., E. Kenchington, E. Zouros et G. C. Rodakis. 2004. «Evidence that the large noncoding sequence is the main control region of maternally and paternally transmitted mitochondrial genomes of the marine mussel (*Mytilus* spp.)». *Genetics*, vol. 167, no 2, p. 835-850.
- Cao, L. Q., B. S. Ort, A. Mizi, G. Pogson, E. Kenchington, E. Zouros et G. C. Rodakis. 2009. «The control region of maternally and paternally inherited mitochondrial genomes of three species of the sea mussel genus *Mytilus*». *Genetics*, vol. 181, no 3, p. 1045-1056.
- Carapelli, A., S. Comandi, P. Convey, F. Nardi et F. Frati. 2008. «The complete mitochondrial genome of the antarctic springtail *Cryptopygus antarcticus* (Hexapoda: Collembola)». *BMC Genomics*, vol. 9, no 315, doi : 10.1186/1471-2164-9-315.

- Carrodeguas, J. A., K. Theis, D. F. Bogenhagen et C. Kisker. 2001. «Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase γ , pol γ b, functions as a homodimer». *Molecular cell*, vol. 7, no 1, p. 43-54.
- Chakrabarti, R., S. Shepardson, M. Karmakar, R. Trdan, J. Walker, R. Shandilya, D. Stewart, S. Vijayaraghavan et W. Hoeh. 2009. «Extra-mitochondrial localization and likely reproductive function of a female-transmitted cytochrome c oxidase subunit II protein». *Development Growth and Differentiation*, vol. 51, no 5, p. 511-519.
- Chakrabarti, R., J. M. Walker, E. G. Chapman, S. P. Shepardson, R. J. Trdan, J. P. Curole, G. T. Watters, D. T. Stewart, S. Vijayaraghavan et W. R. Hoeh. 2007. «Reproductive function for a c-terminus extended, male-transmitted cytochrome c oxidase subunit II protein expressed in both spermatozoa and eggs». *Febs Letters*, vol. 581, no 27, p. 5213-5219.
- Chakrabarti, R., J. M. Walker, D. T. Stewart, R. J. Trdan, S. Vijayaraghavan, J. P. Curole et W. R. Hoeh. 2006. «Presence of a unique male-specific extension of c-terminus to the cytochrome c oxidase subunit II protein coded by the male-transmitted mitochondrial genome of *Venustaconcha ellipsiformis* (Bivalvia: Unionoidea)». *Febs Letters*, vol. 580, no 3, p. 862-866.
- Chan, D. C. 2006. «Mitochondrial fusion and fission in mammals». *Annual Review of Cell and Developmental Biology*, vol. 22, p. 79-99.
- Chan, W. Y., T. L. Lee, S. M. Wu, L. Ruszczka, D. Alba, V. Baxendale et O. M. Rennert. 2006. «Transcriptome analyses of male germ cells with serial analysis of gene expression (sage)». *Molecular and Cellular Endocrinology*, vol. 250, no 1-2, p. 8-19.
- Chapman, E. G., H. Piontkivska, J. M. Walker, D. T. Stewart, J. P. Curole et W. R. Hoeh. 2008. «Extreme primary and secondary protein structure variability in the chimeric male-transmitted cytochrome c oxidase subunit II protein in freshwater mussels: Evidence for an elevated amino acid substitution rate in the face of domain-specific purifying selection». *BMC Evolutionary Biology*, vol. 8, no 165, doi : 10.1186/1471-2148-8-165.

- Charlesworth, D. et S. I. Wright. 2001. «Breeding systems and genome evolution». *Current Opinion in Genetics and Development*, vol. 11, no 6, p. 685-690.
- Chinnery, P. F. 2002. «Modulating heteroplasmy». *Trends in Genetics*, vol. 18, no 4, p. 173-176.
- Clarke, A. H. 1981. *The freshwater molluscs of Canada*. Ottawa (Ontario): National Museums of Canada.
- Clarke, A. H., Jr. et A. M. Rick. 1963. «Supplementary records of unionidacea from Nova Scotia with a discussion of the identity of *Anodonta fragilis* Lamarck». Department of Northern Affairs and National Resources. Ottawa (Ontario), National Museum of Canada Bulletin: 15-27.
- Clayton, D. A. 1982. «Replication of animal mitochondrial DNA». *Cell*, vol. 28, no 4, p. 693-705.
- Clayton, D. A. 1991. «Replication and transcription of vertebrate mitochondrial DNA». *Annual Review of Cell Biology*, vol. 7, no 1, p. 453-478.
- Clement, M., D. Posada et K. A. Crandall. 2000. «Tcs: A computer program to estimate gene genealogies». *Molecular Ecology*, vol. 9, no 10, p. 1657-1659.
- Cogswell, A. T., E. L. R. Kenchington et E. Zouros. 2006. «Segregation of sperm mitochondria in two- and four-cell embryos of the blue mussel *Mytilus edulis*: Implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA». *Genome*, vol. 49, no 7, p. 799-807.
- Cohen, N. G. et P. Lestienne. 2001. *Les maladies mitochondriales*. Paris (France): Editions scientifiques et medicales Elsevier.
- Coller, H. A., K. Khrapko, N. D. Bodyak, E. Nekhaeva, P. Herrero-Jimenez et W. G. Thilly. 2001. «High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection». *Nature Genetics*, vol. 28, no 2, p. 147-150.
- Conant, G. C. et K. H. Wolfe. 2008. «GenomeVx: Simple web-based creation of editable circular chromosome maps». *Bioinformatics*, vol. 24, no 6, p. 861-862.

- Copeland, W. C. 2008. «Inherited mitochondrial diseases of DNA replication». *Annual Review of Medicine*, vol. 59, no 1, p. 131-146.
- Cosmides, L. M. et J. Tooby. 1981. «Cytoplasmic inheritance and intragenomic conflict». *Journal of Theoretical Biology*, vol. 89, no 1, p. 83-129.
- Croteau, D. L. et V. A. Bohr. 1997. «Repair of oxidative damage to nuclear and mitochondrial DNA in mammalian cells». *Journal of Biological Chemistry*, vol. 272, no 41, p. 25409-25412.
- Cummins, J. 1998. «Mitochondrial DNA in mammalian reproduction». *Reviews of Reproduction*, vol. 3, no 3, p. 172-182.
- Curole, J. P. et T. D. Kocher. 2002. «Ancient sex-specific extension of the cytochrome c oxidase II gene in bivalves and the fidelity of doubly-uniparental inheritance». *Molecular Biology and Evolution*, vol. 19, no 8, p. 1323-1328.
- Curole, J. P. et T. D. Kocher. 2005. «Evolution of a unique mitotype-specific protein-coding extension of the cytochrome c oxidase II gene in freshwater mussels (Bivalvia: Unionoida)». *Journal of Molecular Evolution*, vol. 61, no 3, p. 381-389.
- Cyr, F., A. Paquet, A. L. Martel et B. Angers. 2007. «Cryptic lineages and hybridization in freshwater mussels of the genus *Pyganodon* (Unionidae) in northeastern north america». *Canadian Journal of Zoology-Revue Canadienne de Zoologie*, vol. 85, no 12, p. 1216-1227.
- Dalziel, A. C. et D. T. Stewart. 2002. «Tissue-specific expression of male-transmitted mitochondrial DNA and its implications for rates of molecular evolution in *Mytilus* mussels (Bivalvia: Mytilidae)». *Genome*, vol. 45, no 2, p. 348-355.
- Das, J. 2006. «The role of mitochondrial respiration in physiological and evolutionary adaptation». *Bioessays*, vol. 28, no 9, p. 890-901.
- Davis, J. I. et K. C. Nixon. 1992. «Populations, genetic-variation, and the delimitation of phylogenetic species». *Systematic Biology*, vol. 41, no 4, p. 421-435.

- Davison, A. 2006. «The ovotestis: An underdeveloped organ of evolution». *Bioessays*, vol. 28, no 6, p. 642-650.
- Denver, D. R., K. Morris, M. Lynch, L. L. Vassilieva et W. K. Thomas. 2000. «High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*». *Science*, vol. 289, no 5488, p. 2342-2344.
- Devenish, R. J., M. Prescott, X. Roucou et P. Nagley. 2000. «Insights into atp synthase assembly and function through the molecular genetic manipulation of subunits of the yeast mitochondrial enzyme complex». *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, vol. 1458, no 2-3, p. 428-442.
- Doublet, V., C. Souty-Grosset, D. Bouchon, R. Cordaux et I. Marcade. 2008. «A thirty million year-old inherited heteroplasmy». *PLoS ONE*, vol. 3, no 8, p. e2938.
- Doucet-Beaupré, H., S. Breton, E. G. Chapman, P. U. Blier, A. E. Bogan, D. T. Stewart et W. R. Hoeh. 2010. «Mitochondrial phylogenomics of the Bivalvia (Mollusca): Searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA». *BMC Evolutionary Biology*, vol. 10, no 50, doi : 10.1186/1471-2148-10-50.
- Dowling, D. K., U. Friberg et J. Lindell. 2008. «Evolutionary implications of non-neutral mitochondrial genetic variation». *Trends in Ecology and Evolution*, vol. 23, no 10, p. 546-554.
- Downs, J. A. et S. P. Jackson. 2003. «Cancer: Protective packaging for DNA». *Nature*, vol. 424, no 6950, p. 732-734.
- Dowton, M. et A. D. Austin. 1999. «Evolutionary dynamics of a mitochondrial rearrangement "Hot spot" in the Hymenoptera». *Molecular Biology and Evolution*, vol. 16, no 2, p. 298-309.
- Dreyer, H. et G. Steiner. 2006. «The complete sequences and gene organisation of the mitochondrial genomes of the heterodont bivalves *Acanthocardia tuberculata* and *Hiattella arctica* - and the first record for a putative atpase subunit 8 gene in marine bivalves». *Frontiers in Zoology*, vol. 3, no 1, p. 13.
- Edwards, A. W. F. 1972. *Likelihood*. Cambridge: Cambridge University Press.

- Edwards, S. et S. Bensch. 2009. «Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008». *Molecular Ecology*, vol. 18, no 14, p. 2930-2933.
- Ellison, C. K. et R. S. Burton. 2006. «Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*». *Evolution*, vol. 60, no 7, p. 1382-1391.
- Elson, J. L. et R. N. Lightowlers. 2006. «Mitochondrial DNA clonality in the dock: Can surveillance swing the case?». *Trends in Genetics*, vol. 22, no 11, p. 603.
- Evans, B. J. 2008. «Genome evolution and speciation genetics of clawed frogs (*Xenopus* and *Silurana*)». *Frontiers in Bioscience*, vol. 13, p. 4687-4706.
- Everett, E. M., P. J. Williams, G. Gibson et D. T. Stewart. 2004. «Mitochondrial DNA polymorphisms and sperm motility in *Mytilus edulis* (Bivalvia: Mytilidae)». *Journal of Experimental Zoology Part A-Comparative Experimental Biology*, vol. 301A, no 11, p. 906-910.
- Excoffier, L., G. Laval et S. Schneider. 2005. «Arlequin ver. 3.0: An integrated software package for population genetics data analysis ». *Evolutionary Bioinformatics Online*, vol. 1, p. 47-50.
- Excoffier, L., P. E. Smouse et J. M. Quattro. 1992. «Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data». *Genetics*, vol. 131, no 2, p. 479-491.
- Farris, J. S. 1977. «Phylogenetic analysis under dollo's law». *Systematic Zoology*, vol. 26, no 1 p. 77-88
- Felsenstein, J. 1985. «Confidence limits on phylogenies: An approach using the bootstrap». *Evolution*, vol. 39, p. 783-791.
- Fernandez-Silva, P., J. Enriquez et J. Montoya. 2003. «Replication and transcription of mammalian mitochondrial DNA». *Experimental Physiology*, vol. 88, no 1, p. 41-56.

- Fernandez, A. A. et M. R. Morris. 2007. «Sexual selection and trichromatic color vision in primates: Statistical support for the preexisting-bias hypothesis». *The American Naturalist*, vol. 170, no 1, p. 10-20.
- Filipowicz, M., A. Burzynski, B. Smietanka et R. Wenne. 2008. «Recombination in mitochondrial DNA of european mussels *Mytilus*». *Journal of Molecular Evolution*, vol. 67, no 4, p. 377-388.
- Fisher, C. et D. O. F. Skibinski. 1990. «Sex-biased mitochondrial-DNA heteroplasmy in the marine mussel *Mytilus*». *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 242, no 1305, p. 149-156.
- Fisher, T. G. et D. G. Smith. 1994. «Glacial lake agassiz - its northwest maximum extent and outlet in saskatchewan (emerson phase)». *Quaternary Science Reviews*, vol. 13, no 9-10, p. 845-858.
- Folmer, O., M. Black, W. Hoeh, R. Lutz et R. Virjenhoek. 1994. «DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates». *Molecular Marine Biology and Biotechnology*, no 3, p. 294-299.
- Fontaneto, D., E. A. Herniou, T. G. Barraclough, C. Ricci et G. Melone. 2007. «On the reality and recognisability of asexual organisms: Morphological analysis of the masticatory apparatus of bdelloid rotifers». *Zoologica Scripta*, vol. 36, no 4, p. 361-370.
- Force, A., M. Lynch, F. Pickett, A. Amores, Y.-L. Yan et J. Postlethwait. 1999. «Preservation of duplicate genes by complementary, degenerative mutations». *Genetics*, vol. 151, p. 1531 - 1545.
- Francino, M. P. et H. Ochman. 1997. «Strand asymmetries in DNA evolution». *Trends in Genetics*, vol. 13, no 6, p. 240-245.
- Frank, S. A. et L. D. Hurst. 1996. «Mitochondria and male disease». *Nature*, vol. 383, no 6597, p. 224-224.
- Fuller, K. M. et E. Zouros. 1993. «Dispersed discrete length polymorphism of mitochondrial DNA in the scallop *Placopecten magellanicus* (gmelin)». *Current Genetics*, vol. 23, no 4, p. 365-369.

- Garesse, R. et C. G. Vallejo. 2001. «Animal mitochondrial biogenesis and function: A regulatory cross-talk between two genomes». *Gene*, vol. 263, no 1-2, p. 1-16.
- Garrido-Ramos, M. A., D. T. Stewart, B. W. Sutherland et E. Zouros. 1998. «The distribution of male-transmitted and female-transmitted mitochondrial DNA types in somatic tissues of blue mussels: Implications for the operation of doubly uniparental inheritance of mitochondrial DNA». *Genome*, vol. 41, no 6, p. 818-824.
- Gemmell, N. J. et T. L. Braisher. 2001. «Organelle genome evolution». *Trends in Ecology and Evolution*, vol. 16, no 9, p. 489-490.
- Gemmell, N. J., V. J. Metcalf et F. W. Allendorf. 2004. «Mother's curse: The effect of mtDNA on individual fitness and population viability». *Trends in Ecology and Evolution*, vol. 19, no 5, p. 238-244.
- Giles, R. E., H. Blanc, H. M. Cann et D. C. Wallace. 1980. «Maternal inheritance of human mitochondrial DNA». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 77, no 11, p. 6715-6719.
- Giribet, G. 2008. «Bivalvia». In *Phylogeny and evolution of the Mollusca*, W. Ponder et D. R. Lindberg (éd.), p. 105-142. Berkeley (California): University of California Press.
- Giribet, G. 2010. «A new dimension in combining data? The use of morphology and phylogenomic data in metazoan systematics». *Acta Zoologica*, vol. 91, no 1, p. 11-19.
- Giribet, G. et D. L. Distel. 2003. «Bivalve phylogeny and molecular data». In *Molecular systematics and phylogeography of mollusks*, C. Lydeard et D. R. Lindberg (éd.), p. 45-90. Washington, D.C.: Smithsonian Institution Press.
- Giribet, G. et W. Wheeler. 2002. «On bivalve phylogeny: A high-level analysis of the Bivalvia (Mollusca) based on combined morphology and DNA sequence data». *Invertebrate Biology*, vol. 121, no 4, p. 271-324.
- Gissi, C., F. Iannelli et G. Pesole. 2008. «Evolution of the mitochondrial genome of metazoa as exemplified by comparison of congeneric species». *Heredity*, vol. 101, no 4, p. 301-320.

- Gleyzer, N., K. Vercauteren et R. C. Scarpulla. 2005. «Control of mitochondrial transcription specificity factors (tfb1m and tfb2m) by nuclear respiratory factors (nrf-1 and nrf-2) and pgc-1 family coactivators». *Molecular and Cellular Biology*, vol. 25, no 4, p. 1354-1366.
- Graf, D.L. et K.S. Cummings. 2006. «Palaeoheterodont Diversity (Mollusca: Trigonioida + Unionoida): what we know and what we wish we knew about freshwater mussel evolution». *Zoological Journal of the Linnean Society*, vol. 148, p. 343-394.
- Graf, D.L. et K.S. Cummings. 2007. «Review of the systematics and global diversity of freshwater mussel species (Bivalvia: Unionoida) ». *Journal of Mollusca Studies*, vol. 73, p. 291-314.
- Graf, D. L. et D. Ó. Foighil. 2000. «The evolution of brooding characters among the freshwater pearly mussels (Bivalvia: Unionoidea) of North America». *Journal of Molluscan Studies*, vol. 66, no 2, p. 157-170.
- Grande, C., J. Templado et R. Zardoya. 2008. «Evolution of gastropod mitochondrial genome arrangements». *BMC Evolutionary Biology*, vol. 8, no 61, doi : 10.1186/1471-2148-8-61.
- Green, D. R. et J. C. Reed. 1998. «Mitochondria and apoptosis». *Science*, vol. 281, no 5381, p. 1309-1312.
- Gross, N. J., G. S. Getz et M. Rabinowitz. 1969. «Apparent turnover of mitochondrial deoxyribonucleic acid and mitochondrial phospholipids in the tissues of the rat». *Journal of Biological Chemistry*, vol. 244, no 6, p. 1552-1562.
- Grossman, L. I., D. E. Wildman, T. R. Schmidt et M. Goodman. 2004. «Accelerated evolution of the electron transport chain in anthropoid primates». *Trends Genet*, vol. 20, no 11, p. 578 - 585.
- Haag-Liautard, C., N. Coffey, D. Houle, M. Lynch, B. Charlesworth et P. D. Keightley. 2008. «Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*». *PLoS Biology*, vol. 6, no 8, p. e204.

- Haag, W. R. et J. Leann Staton. 2003. «Variation in fecundity and other reproductive traits in freshwater mussels». *Freshwater Biology*, vol. 48, no 12, p. 2118-2130.
- Hajibabaei, M., G. A. C. Singer, P. D. N. Hebert et D. A. Hickey. 2007. «DNA barcoding: How it complements taxonomy, molecular phylogenetics and population genetics». *Trends in Genetics*, vol. 23, no 4, p. 167-172.
- Hall, T. A. 1999. «Bioedit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt. ». *Nucleic Acids Symposium Series*, vol. 41, p. 95-98.
- Hanlon, S. D. et D. G. Smith. 1999. «An attempt to detect *Pyganodon fragilis* (Mollusca: Unionidae) in Maine ». *Northeastern Naturalist*, vol. 6, no 2, p. 119-132.
- Hassanin, A., N. Leger et J. Deutsch. 2005. «Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inferences». *Systematic Biology*, vol. 54, p. 277 - 298.
- Hatefi, Y. 1985. «The mitochondrial electron transport and oxidative phosphorylation system ». *Annual Review of Biochemistry*, vol. 54, p. 1015-1069.
- He, X. et J. Zhang. 2006. «Transcriptional reprogramming and backup between duplicate genes: Is it a genomewide phenomenon?». *Genetics*, vol. 172, no 2, p. 1363-1367.
- Healy, J. M. 1989. «Spermiogenesis and spermatozoa in the relict bivalve genus *neotrigonia*: Relevance to trigonioid relationships, particularly Unionoidea». *Marine Biology*, vol. 103, no 1, p. 75-85.
- Hebert, P. D. N., A. Cywinska, S. L. Ball et J. R. deWaard. 2003. «Biological identifications through DNA barcodes». *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 270, no 1512, p. 313 - 321.
- Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemlak et C. M. Francis. 2004. «Identification of birds through DNA barcodes». *PLoS Biology*, vol. 2, no 10, p. e312.

- Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham, J. B. Johnson, L. Rissler, P. F. Victoriano et A. D. Yoder. 2010. «Phylogeography's past, present, and future: 10 years after». *Molecular Phylogenetics and Evolution*, vol. 54, no 1, p. 291-301.
- Hittinger, C. T. et S. B. Carroll. 2007. «Gene duplication and the adaptive evolution of a classic genetic switch». *Nature*, vol. 449, no 7163, p. 677-681.
- Hoeh, W. R. 1990. «Phylogenetic relationships among eastern North American *Anodonta* (Bivalvia: Unionidae)». *Malacological Review*, vol. 23, p. 63-82.
- Hoeh, W. R., M. B. Black, R. Gustafson, A. E. Bogan, R. A. Lutz et R. C. Vrijenhoek. 1998. «Testing alternative hypotheses of Neotrigonia (Bivalvia: Trigonioidea) phylogenetic relationships using cytochrome c oxidase subunit I DNA sequences». *Malacologia*, vol. 40, no 1-2, p. 267-278.
- Hoeh, W. R., K. H. Blakley et W. M. Brown. 1991. «Heteroplasmy suggests limited biparental inheritance of *Mytilus* mitochondrial-DNA». *Science*, vol. 251, no 5000, p. 1488-1490.
- Hoeh, W. R., A. E. Bogan et W. H. Heard. 2001. «A phylogenetic perspective on the evolution of morphological and reproductive characteristics in the Unionoidea». In *Ecology and evolution of the freshwater mussels unionoidea*, G. Bauer et K. Wächter (éd.), p. 257-280. Berlin: Springer-Verlag.
- Hoeh, W. R., K. S. Frazer, E. Naranjo-Garcia, M. B. Black, D. J. Berg et S. I. Guttman. 1998. «Correlation between mating system and distribution of genetic variation in *Utterbackia* (Bivalvia: Unionidae)». *Journal of Shellfish Research*, vol. 17, no 5, p. 1383-1393.
- Hoeh, W. R., D. T. Stewart et S. I. Guttman. 2002. «High fidelity of mitochondrial genome transmission under the doubly uniparental mode of inheritance in freshwater mussels (Bivalvia:Unionoidea)». *Evolution*, vol. 56, no 11, p. 2252-2261.
- Hoeh, W. R., D. T. Stewart, C. Saavedra, B. W. Sutherland et E. Zouros. 1997. «Phylogenetic evidence for role-reversals of gender-associated

- mitochondrial DNA in *Mytilus* (Bivalvia: Mytilidae)». *Molecular Biology and Evolution*, vol. 14, no 9, p. 959-967.
- Hoeh, W. R., D. T. Stewart, B. W. Sutherland et E. Zouros. 1996a. «Cytochrome c oxidase sequence comparisons suggest an unusually high rate of mitochondrial DNA evolution in *Mytilus* (Mollusca: Bivalvia)». *Molecular Biology and Evolution*, vol. 13, no 2, p. 418-421.
- Hoeh, W. R., D. T. Stewart, B. W. Sutherland et E. Zouros. 1996b. «Multiple origins of gender-associated mitochondrial DNA lineages in bivalves (Mollusca: Bivalvia)». *Evolution*, vol. 50, no 6, p. 2276-2286.
- Hoffmann, R. J., J. L. Boore et W. M. Brown. 1992. «A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*». *Genetics*, vol. 131, no 2, p. 397-412.
- Holt, I. J., H. E. Lorimer et H. T. Jacobs. 2000. «Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA». *Cell*, vol. 100, no 5, p. 515-524.
- Huelsenbeck, J. P. et F. Ronquist. 2001. «Mrbayes: Bayesian inference of phylogenetic trees». *Bioinformatics*, vol. 17, no 8, p. 754-755.
- Huminięcki, L. et K. H. Wolfe. 2004. «Divergence of spatial gene expression profiles following species-specific gene duplications in human and mouse». *Genome Research*, vol. 14, no 10a, p. 1870-1879.
- Hüttemann, M., S. Jaradat et L. I. Grossman. 2003. «Cytochrome c oxidase of mammals contains a testes-specific isoform of subunit VIb—the counterpart to testes-specific cytochrome c?». *Molecular Reproduction and Development*, vol. 66, no 1, p. 8-16.
- Inoue, J. G., M. Miya, K. Tsukamoto et M. Nishida. 2001. «Complete mitochondrial DNA sequence of *Conger myriaster* (Teleostei:Anguilliformes): Novel gene order for vertebrate mitochondrial genomes and the phylogenetic implications for anguilliform families». *Journal of Molecular Evolution*, vol. 52, no 4, p. 311-320.
- Jameson, D., A. P. Gibson, C. Hudelot et P. G. Higgs. 2003. «Ogre: A relational database for comparative analysis of mitochondrial genomes». *Nucleic Acids Research*, vol. 31, p. 202 - 206.

- Janke, A., G. Feldmaierfuchs, W. K. Thomas, A. Vonhaeseler et S. Paabo. 1994. «The marsupial mitochondrial genome and the evolution of placental mammals». *Genetics*, vol. 137, no 1, p. 243-256.
- Jenuth, J. P., A. C. Peterson, K. Fu et E. A. Shoubridge. 1996. «Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA». *Nature Genetics*, vol. 14, no 2, p. 146-151.
- Jha, M., J. Côté, W. R. Hoeh, P. U. Blier et D. T. Stewart. 2008. «Sperm motility in *Mytilus edulis* in relation to mitochondrial DNA polymorphisms: Implications for the evolution of doubly uniparental inheritance in bivalves». *Evolution*, vol. 62, no 1, p. 99-106.
- Jones, D. T., W. R. Taylor et J. M. Thornton. 1992. «The rapid generation of mutation data matrices from protein sequences». *Computer Applications In The Biosciences*, vol. 8, no 3, p. 275-282.
- Jukes, T. H. et S. Osawa. 1990. «The genetic code in mitochondria and chloroplasts». *Cellular and Molecular Life Sciences*, vol. 46, no 11, p. 1117-1126.
- Kagawa, W., H. Kurumizaka, S. Ikawa, S. Yokoyama et T. Shibata. 2001. «Homologous pairing promoted by the human rad52 protein». *Journal of Biological Chemistry*, vol. 276, no 37, p. 35201-35208.
- Kai, Y., C. Takamatsu, K. Tokuda, M. Okamoto, K. Irita et S. Takahashi. 2006. «Rapid and random turnover of mitochondrial DNA in rat hepatocytes of primary culture». *Mitochondrion*, vol. 6, no 6, p. 299-304.
- Kat, P. W. 1983. «Genetic and morphological divergence among nominal species of North American *Anodonta* (Bivalvia, Unionidae)». *Malacologia*, vol. 23, no 1, p. 361-374.
- Kat, P. W. 1986. «Hybridization in a unionid faunal suture zone». *Malacologia*, vol. 27, p. 107-125.
- Kayal, E., B. Bentlage, A.G. Collins, M. Kayal, S. Pirro et D.V. Lavrov. 2012. «Evolution of linear mitochondrial genomes in medusozoan cnidarians». *Genome Biology and Evolution*, vol 4, no 1, p.1-12.

- Kenchington, E., B. MacDonald, L. Q. Cao, D. Tsagkarakis et E. Zouros. 2002. «Genetics of mother-dependent sex ratio in blue mussels (*Mytilus* spp.) and implications for doubly uniparental inheritance of mitochondrial DNA». *Genetics*, vol. 161, no 4, p. 1579-1588.
- Kenchington, E. L., L. Hamilton, A. Cogswell et E. Zouros. 2009. «Paternal mtDNA and maleness are co-inherited but not causally linked in Mytilid mussels». *PLoS One*, vol. 4, no 9, p. e6976.
- Kenyon, L. et C. T. Moraes. 1997. «Expanding the functional human mitochondrial DNA database by the establishment of primate xenomitochondrial cybrids». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no 17, p. 9131-9135.
- Kiyomoto, M., A. Komaru, J. Scarpa, K. T. Wada, E. Danton et M. Awaji. 1996. «Abnormal gametogenesis, male dominant sex ratio, and sertoli cell morphology in induced triploid mussels, *Mytilus galloprovincialis*». *Zoological Science*, vol. 13, no 3, p. 393-402.
- Knudsen, B., A. B. Kohn, B. Nahir, C. S. McFadden et L. L. Moroz. 2006. «Complete DNA sequence of the mitochondrial genome of the sea slug, *Aplysia californica*: Conservation of the gene order in Euthyneura». *Molecular Phylogenetics and Evolution*, vol. 38, no 2, p. 459.
- Koepfli, K.-P., K. Deere, G. Slater, C. Begg, K. Begg, L. Grassman, M. Lucherini, G. Veron et R. Wayne. 2008. «Multigene phylogeny of the mustelidae: Resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation». *BMC Biology*, vol. 6, no 10, doi : 10.1186/1741-7007-6-10.
- Korhonen, J. A., X. H. Pham, M. Pellegrini et M. Falkenberg. 2004. «Reconstitution of a minimal mtDNA replisome *in vitro*». *EMBO Journal*, vol. 23, no 12, p. 2423-2429.
- Krauss, S., C.-Y. Zhang et B. B. Lowell. 2005. «The mitochondrial uncoupling-protein homologues». *Nature Reviews Molecular Cell Biology*, vol. 6, no 3, p. 248-261.

- Krebs, R. A. 2004. «Combining paternally and maternally inherited mitochondrial DNA for analysis of population structure in mussels». *Molecular Ecology*, vol. 13, no 6, p. 1701-1705.
- Kujoth, G. C., A. Hiona, T. D. Pugh, S. Someya, K. Panzer, S. E. Wohlgemuth, T. Hofer, A. Y. Seo, R. Sullivan, W. A. Jobling, J. D. Morrow, H. Van Remmen, J. M. Sedivy, T. Yamasoba, M. Tanokura, R. Weindruch, C. Leeuwenburgh et T. A. Prolla. 2005. «Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging». *Science*, vol. 309, no 5733, p. 481-484.
- Kumar, S., K. Tamura, I. B. Jakobsen et M. Nei. 2001. «Mega2: Molecular evolutionary genetics analysis software». *Bioinformatics*, vol. 17, no 12, p. 1244-1245.
- Kurabayashi, A., M. Sumida, H. Yonekawa, F. Glaw, M. Vences et M. Hasegawa. 2008. «Phylogeny, recombination, and mechanisms of stepwise mitochondrial genome reorganization in mantellid frogs from Madagascar». *Molecular Biology and Evolution*, vol. 25, no 5, p. 874-891.
- Kuroiwa, T. 2010a. «Mechanisms of organelle division and inheritance and their implications regarding the origin of eukaryotic cells». *Proceedings of the Japan Academy Series B-Physical and Biological Sciences*, vol. 86, no 5, p. 455-471.
- Kuroiwa, T. 2010b. «Review of cytological studies on cellular and molecular mechanisms of uniparental (maternal or paternal) inheritance of plastid and mitochondrial genomes induced by active digestion of organelle nuclei (nucleoids)». *Journal of Plant Research*, vol. 123, no 2, p. 207-230.
- Kuroiwa, T., H. Kuroiwa, A. Sakai, H. Takahashi, K. Toda et R. Itoh. 1998. «The division apparatus of plastids and mitochondria». *International review of cytology*, vol. 181, p. 1-41.
- Ladoukakis, E. D. et E. Zouros. 2001. «Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA». *Molecular Biology and Evolution*, vol. 18, no 7, p. 1168-1175.
- Ladoukakis, E. D., Theologidis, I., Rodakis, G. C. et E. Zouros. 2012. «Homologous recombination between highly diverged mitochondrial

- sequences: examples from maternally and paternally transmitted genomes». *Molecular Biology and Evolution*, vol. 28, no 6, p. 1847-1859.
- Lane, N. et W. Martin. 2010. «The energetics of genome complexity». *Nature*, vol. 467, p. 929-934.
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson et D. G. Higgins. 2007. «Clustal W and clustal X version 2.0». *Bioinformatics*, vol. 23, no 21, p. 2947-2948.
- Laslett, D. et B. Canback. 2008. «Arwen: A program to detect tRNA genes in metazoan mitochondrial nucleotide sequences». *Bioinformatics*, vol. 24, no 2, p. 172-175.
- Le Quesne, W. J. 1974. «The uniquely evolved character concept and its cladistic application». *Systematic Zoology*, vol. 23, no 4, p. 513-517
- Le, T. H., D. Blair, T. Agatsuma, P.-F. Humair, N. J. H. Campbell, M. Iwagami, D. T. J. Littlewood, B. Peacock, D. A. Johnston, J. Bartley, D. Rollinson, E. A. Herniou, D. S. Zarlenga et D. P. McManus. 2000. «Phylogenies inferred from mitochondrial gene orders — a cautionary tale from the parasitic flatworms». *Molecular Biology and Evolution*, vol. 17, no 7, p. 1123-1125.
- Lee, W. J. et T. D. Kocher. 1995. «Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome - early establishment of the vertebrate genome organization». *Genetics*, vol. 139, no 2, p. 873-887.
- Leitao, A. et R. Chaves. 2008. «Banding for chromosomal identification in bivalves: A 20-year history». *Aquaculture 1. Dynamic Biochemistry, Process Biotechnology and Molecular Biology*, vol. 2, Special Issue 1, p. 44 - 49.
- Liu, H. P., J. B. Mitton et S. K. Wu. 1996. «Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the freshwater mussel, *Anodonta grandis grandis*». *Evolution*, vol. 50, no 2, p. 952-957.

- Longley, M. J., D. Nguyen, T. A. Kunkel et W. C. Copeland. 2001. «The fidelity of human DNA polymerase γ with and without exonucleolytic proofreading and the p55 accessory subunit». *Journal of Biological Chemistry*, vol. 276, no 42, p. 38555-38562.
- Longo, F. et E. Anderson. 1969 «Cytological aspects of fertilization in the lamellibranch, *Mytilus edulis*. II. Development of the male pronucleus and the association of the maternally and paternally derived chromosomes». *Journal of Experimental Zoology* vol. 172, no 1, p. 97-119.
- Loublier, S., M. Schiff, P. Benit et P. Rustin. 2009. «Mitochondrial diseases: A particular medicine?». *Immuno-analyse et Biologie Spécialisée*, vol. 24, no 5-6, p. 240-253.
- Lowe, T. M. et S. R. Eddy. 1997. «tRNAscan-se: A program for improved detection of transfer RNA genes in genomic sequence». *Nucleic Acids Research*, vol. 25, no 5, p. 955-964.
- Lydeard, C., R. H. Cowie, W. F. Ponder, A. E. Bogan, P. Bouchet, S. A. Clark, K. S. Cummings, T. J. Frest, O. Gargominy, D. G. Herbert, R. Hershler, K. E. Perez, B. Roth, M. Seddon, E. E. Strong et F. G. Thompson. 2004. «The global decline of nonmarine mollusks». *Bioscience*, vol. 54, no 4, p. 321-330.
- Lynch, M. 1997. «Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes». *Molecular Biology and Evolution*, vol. 14, no 9, p. 914-925.
- Lynch, M. et J. S. Conery. 2000. «The evolutionary fate and consequences of duplicate genes». *Science*, vol. 290, no 5494, p. 1151-1155.
- Lynch, M., W. Sung, K. Morris, N. Coffey, C. R. Landry, E. B. Dopman, W. J. Dickinson, K. Okamoto, S. Kulkarni, D. L. Hartl, et W. K. Thomas. 2008. «A genome-wide view of the spectrum of spontaneous mutations in yeast». *Proceedings of the National Academy of Sciences*, vol. 105, no 27, p. 9272-9277.
- Lynch, M. 2010. « Evolution of the mutation rate». *Trends in Genetics*. vol. 26, no 8, p.345-352.

- Macey, J. R., A. Larson, N. B. Ananjeva, Z. Fang et T. J. Papenfuss. 1997. «Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome». *Molecular Biology and Evolution*, vol. 14, no 1, p. 91-104.
- Maddison, D. R. et W. P. Maddison. 2003. *Macclade 4: Analysis of parsimony and character evolution*. 4.06th edition. Sunderland: Sinauer Associates.
- Maddison, W. P. et D. R. Maddison. 2008. «Mesquite: A modular system for evolutionary analysis.». Version 2.6. [<http://mesquiteproject.org>]
- Magnus, G. et J. Keizer. 1998. «Model of beta -cell mitochondrial calcium handling and electrical activity. II. Mitochondrial variables». *Am J Physiol Cell Physiol*, vol. 274, no 4, p. C1174-1184.
- Margulis, L. 1981. *Symbiosis in cell evolution: Life and its environment on the early earth*. San Francisco (California): W. H. Freeman
- Marshall, D. C., C. Simon et T. R. Buckley. 2006. «Accurate branch length estimation in partitioned bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors». *Systematic Biology*, vol. 55, no 6, p. 993-1003.
- Martinou, J. C., S. Desagher et B. Antonsson. 2000. «Cytochrome c release from mitochondria: All or nothing». *Nature Cell Biology*, vol. 2, no 3, p. E41-E43.
- May-Panloup, P., M. F. Chrétien, Y. Malthiery et P. Reynier. 2006. «ADN mitochondrial du spermatozoïde». *Gynécologie Obstétrique et Fertilité*, vol. 34, no 9, p. 847-854.
- McDonald, J. et M. Kreitman. 1991. «Adaptive protein evolution at the Adh locus in *Drosophila*». *Nature*, vol. 351, p. 652-654.
- McKenzie, M., M. Chiotis, C. A. Pinkert et I. A. Trounce. 2003. «Functional respiratory chain analyses in murid xenomitochondrial cybrids expose coevolutionary constraints of cytochrome b and nuclear subunits of complex III». *Molecular Biology and Evolution*, vol. 20, no 7, p. 1117-1124.

- McKenzie, M. et I. Trounce. 2000. «Expression of *Rattus norvegicus* mtDNA in *Mus musculus* cells results in multiple respiratory chain defects». *Journal of Biological Chemistry*, vol. 275, no 40, p. 31514-31519.
- McMahon, R. 1991. «Mollusca: Bivalvia.». In *Ecology and classification of north american freshwater invertebrates.*, J. H. Thorp et A. P. Covich (éd.). p. 315-399. San Diego, USA: Academic Press.
- Meiklejohn, C. D., K. L. Montooth et D. M. Rand. 2007. «Positive and negative selection on the mitochondrial genome». *Trends In Genetics*, vol. 23, no 6, p. 259-263.
- Michelle, E. G., M. Francis, R. L. Sharon, L. C. Catherine, F. H. Jennifer, M. Anne, R. Franklin et L. W. Steven. 2001. «Quantification of mitochondrial DNA in peripheral blood mononuclear cells and subcutaneous fat using real-time polymerase chain reaction». *Journal of clinical virology:the official publication of the Pan American Society for Clinical Virology*, vol. 22, no 3, p. 241-247.
- Milbury, C. A. et P. M. Gaffney. 2005. «Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*». *Marine Biotechnology*, vol. 7, p. 697 - 712.
- Mindell, D. P., M. D. Sorenson et D. E. Dimcheff. 1998. «Multiple independent origins of mitochondrial gene order in birds». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no 18, p. 10693-10697.
- Mishmar, D., E. Ruiz-Pesini, P. Golik, V. Macaulay, A. G. Clark, S. Hosseini, M. Brandon, K. Easley, E. Chen, M. D. Brown, R. I. Sukernik, A. Olckers et D. C. Wallace. 2003. «Natural selection shaped regional mtDNA variation in humans». *Proceedings of the National Academy of Sciences*, vol. 100, no 1, p. 171 - 176.
- Mizi, A., E. Zouros, N. Moschonas et G. C. Rodakis. 2005. «The complete maternal and paternal mitochondrial genomes of the mediterranean mussel *Mytilus galloprovincialis*: Implications for the doubly uniparental inheritance mode of mtDNA». *Molecular Biology and Evolution*, vol. 22, no 4, p. 952-967.
- Mock, K. E., J. C. Brim-Box, M. P. Miller, M. E. Downing et W. R. Hoeh. 2004. «Genetic diversity and divergence among freshwater mussel

- (*Anodonta*) populations in the Bonneville Basin of Utah». *Molecular Ecology*, vol. 13, no 5, p. 1085-1098.
- Montooth, K. L. et D. M. Rand. 2008. «The spectrum of mitochondrial mutation differs across species». *PLoS Biology*, vol. 6, no 8, p. e213.
- Moore, R. C. et M. D. Purugganan. 2003. «The early stages of duplicate gene evolution». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no 26, p. 15682-15687.
- Moritz, C., T. E. Dowling et W. M. Brown. 1987. «Evolution of animal mitochondrial DNA: Relevance for population biology and systematics». *Annual Review of Ecology, Evolution, and Systematics*, vol. 18, p. 269 - 292.
- Mueller, R. L. et J. L. Boore. 2005. «Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders». *Molecular Biology and Evolution*, vol. 22, no 10, p. 2104-2112.
- Murphy, N. P., D. Carey, L. R. Castro, M. Dowton et A. D. Austin. 2007. «Phylogeny of the platygastroid wasps (hymenoptera) based on sequences from the 18s rRNA, 28s rRNA and cytochrome oxidase I genes: Implications for the evolution of the ovipositor system and host relationships». *Biological Journal of the Linnean Society*, vol. 91, no 4, p. 653-669.
- Nagel. 2000. «Testing hypotheses on the dispersal and evolutionary history of freshwater mussels (Mollusca: Bivalvia: Unionidae)». *Journal of Evolutionary Biology*, vol. 13, no 5, p. 854-865.
- Nakada, K., K. Inoue, T. Ono, K. Isobe, A. Ogura, Y.-I. Goto, I. Nonaka et J.-I. Hayashi. 2001. «Inter-mitochondrial complementation: Mitochondria-specific system preventing mice from expression of disease phenotypes by mutant mtDNA». *Nature Medicine*, vol. 7, no 8, p. 934-940.
- Nakada, K., A. Sato, K. Yoshida, T. Morita, H. Tanaka, S. I. Inoue, H. Yonekawa et J. I. Hayashi. 2006. «Mitochondria-related male infertility». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no 41, p. 15148-15153.

- Nalepa, T. F. 1994. «Decline of native unionid bivalves in lake st clair after infestation by the zebra mussel, *Dreissena polymorpha*». *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 51, no 10, p. 2227-2233.
- Nass, M. M. K. 1969. «Mitochondrial DNA i. Intramitochondrial distribution and structural relations of single-and double-length circular DNA. ». *Journal of Molecular Evolution*, vol. 42, p. 521-528.
- Nei, M. et T. Gojobori. 1986. «Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions». *Molecular Biology and Evolution*, vol. 3, no 5, p. 418-426.
- Neiman, M. et D. R. Taylor. 2009. «The causes of mutation accumulation in mitochondrial genomes». *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 276, no 1660, p. 1201-1209.
- Nishimura, Y., T. Yoshinari, K. Naruse, T. Yamada, K. Sumi, H. Mitani, T. Higashiyama et T. Kuroiwa. 2006. «Active digestion of sperm mitochondrial DNA in single living sperm revealed by optical tweezers». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no 5, p. 1382-1387.
- Nowak, M. A., M. C. Boerlijst, J. Cooke et J. M. Smith. 1997. «Evolution of genetic redundancy». *Nature*, vol. 388, no 6638, p. 167-171.
- Obata, M., C. Kamiya, K. Kawamura et A. Komaru. 2006. «Sperm mitochondrial DNA transmission to both male and female offspring in the blue mussel *Mytilus galloprovincialis*». *Development Growth and Differentiation*, vol. 48, no 4, p. 253-261.
- Obata, M. et A. Komaru. 2005. «Specific location of sperm mitochondria in mussel *Mytilus galloprovincialis* zygotes stained by mitotracker». *Development Growth and Differentiation*, vol. 47, no 4, p. 255-263.
- Obata, M., N. Sano, K. Kawamura et A. Komaru. 2007. «Inheritance of two mt type mitochondrial DNA from sperm and unfertilized eggs to offspring in *Mytilus galloprovincialis*». *Development Growth and Differentiation*, vol. 49, no 4, p. 335-344.
- Obata, M., M. Shimizu, N. Sano et A. Komaru. 2008. «Maternal inheritance of mitochondrial DNA (mtDNA) in the pacific oyster (*Crassostrea gigas*):

A preliminary study using mtDNA sequence analysis with evidence of random distribution of mitotracker-stained sperm mitochondria in fertilized eggs». *Zoological Science*, vol. 25, no 3, p. 248-254.

Ohno, S. 1970. *Evolution by gene duplication*. New York, New York: Springer-Verlag.

Okamoto, K. et J. M. Shaw. 2005. «Mitochondrial morphology and dynamics in yeast and multicellular eukaryotes». *Annual Review of Genetics*, vol. 39, no 1, p. 503-536.

Okimoto, R., J. L. Macfarlane, D. O. Clary et D. R. Wolstenholme. 1992. «The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*». *Genetics*, vol. 130, p. 471 - 498.

Ono, T., K. Isobe, K. Nakada et J.-I. Hayashi. 2001. «Human cells are protected from mitochondrial dysfunction by complementation of DNA products in fused mitochondria». *Nature Genetics*, vol. 28, no 3, p. 272-275.

Ort, B. S. et G. H. Pogson. 2007. «Molecular population genetics of the male and female mitochondrial DNA molecules of the california sea mussel, *Mytilus californianus*». *Genetics*, vol. 177, no 2, p. 1087-1099.

Pagel, M. 1999. «The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies». *Systematic Biology*, vol. 48, no 3, p. 612-622.

Palumbi, S. R. 1996. «Nucleic acids II: The polymerase chain reaction». In *Molecular systematics*, D. M. Hillis, C. Moritz et B. K. Mable (éd.), p. 205-247. Sunderland: Sinauer Associates.

Paquet, A., I. Picard, F. Caron et S. Roux. 2004. «Les mulettes au québec ». *Le Naturaliste canadien*, vol. 129, no 1, p. 78-85.

Parekh, A. B. 2003. «Mitochondrial regulation of intracellular Ca^{2+} signaling: More than just simple Ca^{2+} buffers». *News in Physiological Sciences*, vol. 18, no 6, p. 252-256.

Parsons, T. J., D. S. Muniec, K. Sullivan, N. Woodyatt, R. Alliston-Greiner, M. R. Wilson, D. L. Berry, K. A. Holland, V. W. Weedn, P. Gill et M. M. Holland. 1997. «A high observed substitution rate in the human

- mitochondrial DNA control region». *Nature Genetics*, vol. 15, no 4, p. 363-368.
- Passamonti, M. 2007. «An unusual case of gender-associated mitochondrial DNA heteroplasmy: The mytilid *Musculista senhousia* (Mollusca Bivalvia)». *BMC Evolutionary Biology*, vol. 7 (S2), no S7, doi : 10.1186/1471-2148-7-S2-S7.
- Passamonti, M., J. L. Boore et V. Scali. 2003. «Molecular evolution and recombination in gender-associated mitochondrial DNAs of the Manila clam *Tapes philippinarum*». *Genetics*, vol. 164, no 2, p. 603-611.
- Passamonti, M. et F. Ghiselli. 2009. «Doubly uniparental inheritance: Two mitochondrial genomes, one precious model for organelle DNA inheritance and evolution». *DNA and Cell Biology*, vol. 28, no 2, p. 79-89.
- Passamonti, M. et V. Scali. 2001. «Gender-associated mitochondrial DNA heteroplasmy in the venerid clam *Tapes philippinarum* (Mollusca Bivalvia)». *Current Genetics*, vol. 39, no 2, p. 117-124.
- Paz, M., A. Mikhailov et M. Torrado. 2001. «Sexual differentiation of the somatic gonad tissue in marine bivalve mollusks: Esterase- and fibronectin-like recognition signals». *The international Journal of Developmental Biology*, vol. 45, p. S119-S120.
- Paz, M., M. Torrado, L. I. Korochkin et A. T. Mikhailov. 2005. «Esterase-like and fibronectin-like polypeptides share similar sex-cell-biased patterns in the gonad of hermaphroditic and gonochoric species of bivalve mollusks». *Cell and Tissue Research*, vol. 322, no 3, p. 475-489.
- Perna, N. T. et T. D. Kocher. 1995. «Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes». *Journal of Molecular Evolution*, vol. 41, no 3, p. 353-358.
- Pesole, G., C. Gissi, A. De Chirico et C. Saccone. 1999. «Nucleotide substitution rate of mammalian mitochondrial genomes». *Journal of Molecular Evolution*, vol. 48, no 4, p. 427-434.
- Pons, J., T. G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. P. Duran, S. Hazell, S. Kamoun, W. D. Sumlin et A. P. Vogler. 2006. «Sequence-

- based species delimitation for the DNA taxonomy of undescribed insects». *Systematic Biology*, vol. 55, no 4, p. 595-609.
- Poulton, J. 1995. «Transmission of mtDNA: Cracks in the bottleneck». *American Journal of Human Genetics*, vol. 57, no 2, p. 224-226.
- Quesada, H., H. Stuckas et D. O. F. Skibinski. 2003. «Heteroplasmy suggests paternal co-transmission of multiple genomes and pervasive reversion of maternally into paternally transmitted genomes of mussel (*Mytilus*) mitochondrial DNA». *Journal of Molecular Evolution*, vol. 57, p. S138-S147.
- Quesada, H., R. Wenne et D. O. F. Skibinski. 1999. «Interspecies transfer of female mitochondrial DNA is coupled with role-reversals and departure from neutrality in the mussel *Mytilus trossulus*». *Molecular Biology and Evolution*, vol. 16, no 5, p. 655-665.
- Rand, D. M. 2001. «The units of selection on mitochondrial DNA». *Annual Review of Ecology and Systematics*, vol. 32, p. 415-448.
- Rand, D. M. 2008. «Mitigating mutational meltdown in mammalian mitochondria». *PLoS Biology*, vol. 6, no 2, p. e35.
- Rand, D. M., A. G. Clark et L. M. Kann. 2001. «Sexually antagonistic cytonuclear fitness interactions in *Drosophila melanogaster*». *Genetics*, vol. 159, no 1, p. 173-187.
- Rand, D. M., R. A. Haney et A. J. Fry. 2004. «Cytonuclear coevolution: The genomics of cooperation». *Trends in Ecology and Evolution*, vol. 19, no 12, p. 645-653.
- Rawson, P. D. 2005. «Nonhomologous recombination between the large unassigned region of the male and female mitochondrial genomes in the mussel, *Mytilus trossulus*». *Journal of Molecular Evolution*, vol. 61, no 6, p. 717-732.
- Rawson, P. D. et R. S. Burton. 2006. «Molecular evolution at the cytochrome oxidase subunit 2 gene among divergent populations of the intertidal copepod, *Tigriopus californicus*». *Journal of Molecular Evolution*, vol. 62, no 6, p. 753-764.

- Rawson, P. D. et T. J. Hilbish. 1995. «Evolutionary relationships among the male and female mitochondrial-DNA lineages in the *Mytilus edulis* species complex». *Molecular Biology and Evolution*, vol. 12, no 5, p. 893-901.
- Reyes, A., C. Gissi, G. Pesole et C. Saccone. 1998. «Asymmetrical directional mutation pressure in the mitochondrial genome of mammals». *Molecular Biology and Evolution*, vol. 15, no 8, p. 957-966.
- Ricciardi, A., Nevesm R.J., Rasmussen, J.B. 1998. «Impending extinctions of North America freshwater mussels (Unionoida) following the zebra mussels (*Dreissena polymorpha*) invasion». *Journal of Animal Ecology*, vol. 67, p. 613-619.
- Rigaa, A., M. Monnerot et D. Sellos. 1995. «Molecular cloning and complete nucleotide sequence of the repeated unit and flanking gene of the scallop *Pecten maximus* mitochondrial DNA: putative replication origin features». *Journal of Molecular Evolution*, vol. 41, no 2, p. 189-195.
- Riginos, C., M. J. Hickerson, C. M. Henzler et C. W. Cunningham. 2004. «Differential patterns of male and female mtDNA exchange across the atlantic ocean in the blue mussel, *Mytilus edulis*». *Evolution*, vol. 58, no 11, p. 2438-2451.
- Rodriguez, F., J. L. Oliver, A. Marin et J. R. Medina. 1990. «The general stochastic model of nucleotide substitution». *Journal of Theoretical Biology*, vol. 142, no 4, p. 485-501.
- Rokas, A., E. Ladoukakis et E. Zouros. 2003. «Animal mitochondrial DNA recombination revisited». *Trends in Ecology and Evolution*, vol. 18, no 8, p. 411-417.
- Ronquist, F., J. Huelsenbeck et P. van der Mark (2005). Mrbayes 3.1 manual. [http://mr bayes.Csit.Fsu.Edu/mb3.1_manual.Pdf].
- Ronquist, F. et J. P. Huelsenbeck. 2003. «Mrbayes 3: Bayesian phylogenetic inference under mixed models». *Bioinformatics*, vol. 19, p. 1572 - 1574.
- Roth, C., S. Rastogi, L. Arvestad, K. Dittmar, S. Light, D. Ekman et D. A. Liberles. 2007. «Evolution after gene duplication: Models, mechanisms, sequences, systems, and organisms». *Journal of*

Experimental Zoology Part B: Molecular and Developmental Evolution, vol. 308B, no 1, p. 58-73.

- Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer et R. Rozas. 2003. «DnaSP, DNA polymorphism analyses by the coalescent and other methods». *Bioinformatics*, vol. 19, no 18, p. 2496-2497.
- Ruiz-Pesini, E., A. C. Lapena, C. Diez-Sanchez, A. Perez-Martos, J. Montoya, E. Alvarez, M. Diaz, A. Urrieis, L. Montoro, M. J. Lopez-Perez et J. A. Enriquez. 2000. «Human mtDNA haplogroups associated with high or reduced spermatozoa motility». *American Journal of Human Genetics*, vol. 67, no 3, p. 682-696.
- Ruiz-Pesini, E., D. Mishmar, M. Brandon, V. Procaccio et D. C. Wallace. 2004. «Effects of purifying and adaptive selection on regional variation in human mtDNA». *Science*, vol. 303, no 5655, p. 223-226.
- Saavedra, C., M. I. Reyero et E. Zouros. 1997. «Male-dependent doubly uniparental inheritance of mitochondrial DNA and female-dependent sex-ratio in the mussel *Mytilus galloprovincialis*». *Genetics*, vol. 145, no 4, p. 1073-1082.
- Sagan, L. 1967. «On the origin of mitosing cells». *Journal of Theoretical Biology*, vol. 14, no 3, p. 255-274.
- Sanderson, M. J. 2003. «R8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock». *Bioinformatics*, vol. 19, no 2, p. 301-302.
- Sano, N., M. Obata et A. Komaru. 2007. «Quantitation of the male and female types of mitochondrial DNA in a blue mussel, *Mytilus galloprovincialis*, using real-time polymerase chain reaction assay». *Development Growth and Differentiation*, vol. 49, no 1, p. 67-72.
- Saraste, M. 1999. «Oxidative phosphorylation at the fin de siècle». *Science*, vol. 283, no 5407, p. 1488 - 1493.
- Satoh, M. et T. Kuroiwa. 1991. «Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell ». *Experimental Cell Research*, vol. 196, p. 137-140.

- Scannell, D. R., K. P. Byrne, J. L. Gordon, S. Wong et K. H. Wolfe. 2006. «Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts». *Nature*, vol. 440, no 7082, p. 341-345.
- Scarpulla, R. C. 1997. «Nuclear control of respiratory chain expression in mammalian cells». *Journal of Bioenergetics and Biomembranes*, vol. 29, no 2, p. 109-119.
- Schloesser, D. W. et T. F. Nalepa. 1994. «Dramatic decline of unionid bivalves in offshore waters of western Lake Erie after infestation by the zebra mussel, *Dreissena polymorpha*». *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 51, no 10, p. 2234-2242.
- Schmidt, B., J. McCracken et S. Ferguson-Miller. 2003. «A discrete water exit pathway in the membrane protein cytochrome c oxidase». *Proceedings of the National Academy of Sciences*, vol. 100, no 26, p. 15539 - 15542.
- Selosse, M. A., B. R. Albert et B. Godelle. 2001. «Reducing the genome size of organelles favours gene transfer to the nucleus». *Trends in Ecology and Evolution*, vol. 16, no 3, p. 135-141.
- Serb, J. M. et C. Lydeard. 2003. «Complete mtDNA sequence of the north american freshwater mussel, *Lampsilis ornata* (Unionidae): An examination of the evolution and phylogenetic utility of mitochondrial genome organization in Bivalvia (Mollusca)». *Molecular Biology and Evolution*, vol. 20, no 11, p. 1854-1866.
- Shadel, G. S. 2008. «Expression and maintenance of mitochondrial DNA: New insights into human disease pathology». *American Journal of Pathology*, vol. 172, no 6, p. 1445-1456.
- Shadel, G. S. et D. A. Clayton. 1997. «Mitochondrial DNA maintenance in vertebrates». *Annual Review of Biochemistry*, vol. 66, no 1, p. 409-435.
- Shibata, T. et F. Ling. 2007. «DNA recombination protein-dependent mechanism of homoplasmy and its proposed functions». *Mitochondrion*, vol. 7, no 1-2, p. 17-23.
- Shitara, H., H. Kaneda, A. Sato, K. Inoue, A. Ogura, H. Yonekawa et J.-I. Hayashi. 2000. «Selective and continuous elimination of mitochondria

- microinjected into mouse eggs from spermatids, but not from liver cells, occurs throughout embryogenesis». *Genetics*, vol. 156, no 3, p. 1277-1284.
- Shoubridge, E. A. et T. Wai. 2008. «Medicine: Sidestepping mutational meltdown». *Science*, vol. 319, no 5865, p. 914-915.
- Sites, J. W. et J. C. Marshall. 2004. «Operational criteria for delimiting species». *Annual Review of Ecology Evolution and Systematics*, vol. 35, p. 199-227.
- Skibinski, D. O. F., C. Gallagher et C. M. Beynon. 1994a. «Mitochondrial-DNA inheritance». *Nature*, vol. 368, no 6474, p. 817-818.
- Skibinski, D. O. F., C. Gallagher et C. M. Beynon. 1994b. «Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*». *Genetics*, vol. 138, no 3, p. 801-809.
- Skulachev, V. P. 1999. «Mitochondrial physiology and pathology; concepts of programmed death of organelles, cells and organisms». *Molecular Aspects of Medicine*, vol. 20, no 3, p. 139-184.
- Śmietanka, B., A. Burzynski et R. Wenne. 2009. «Molecular population genetics of male and female mitochondrial genomes in European mussels *Mytilus*». *Marine Biology*, vol. 156, no 5, p. 913-925.
- Smith, D. et M. Snyder. 2007. «Complete mitochondrial DNA sequence of the scallop *Placopecten magellanicus*: Evidence of transposition leading to an uncharacteristically large mitochondrial genome». *Journal of Molecular Evolution*, vol. 65, no 4, p. 380-391.
- Spelbrink, J. N., F.-Y. Li, V. Tiranti, K. Nikali, Q.-P. Yuan, M. Tariq, S. Wanrooij, N. Garrido, G. Comi, L. Morandi, L. Santoro, A. Toscano, G.-M. Fabrizi, H. Somer, R. Croxen, D. Beeson, J. Poulton, A. Suomalainen, H. T. Jacobs, M. Zeviani et C. Larsson. 2001. «Human mitochondrial DNA deletions associated with mutations in the gene encoding twinkie, a phage t7 gene 4-like protein localized in mitochondria». *Nature Genetics*, vol. 28, no 3, p. 223-231.
- Spooner, D. M. 2009. «DNA barcoding will frequently fail in complicated groups: An example in wild potatoes». *American Journal of Botany*, vol. 96, no 6, p. 1177-1189.

- St John, J., D. Sakkas, K. Dimitriadi, A. Barnes, V. Maclin, J. Ramey, C. Barratt et C. De Jonge. 2000. «Failure of elimination of paternal mitochondrial DNA in abnormal embryos». *Lancet*, vol. 355, no 9199, p. 200-200.
- Stewart, D. T., S. Breton, P. U. Blier et W. R. Hoeh. 2009. «Masculinization events and doubly uniparental inheritance of mitochondrial DNA: A model for understanding the evolutionary dynamics of gender-associated mtDNA in mussels». In *Evolutionary biology: Concept, modeling, and application*, P. Pontarotti (éd.), p. 163-173. Berlin: Springer-Verlag Berlin.
- Stewart, D. T., E. R. Kenchington, R. K. Singh et E. Zouros. 1996. «Degree of selective constraint as an explanation of the different rates of evolution of gender-specific mitochondrial DNA lineages in the mussel *Mytilus*». *Genetics*, vol. 143, no 3, p. 1349-1357.
- Stewart, D. T., C. Saavedra, R. R. Stanwood, A. O. Ball et E. Zouros. 1995. «Male and female mitochondrial-DNA lineages in the blue mussel (*Mytilus edulis*) species group». *Molecular Biology and Evolution*, vol. 12, no 5, p. 735-747.
- Stewart, J. B., C. Freyer, J. L. Elson et N. G. Larsson. 2008a. «Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease». *Nature Reviews Genetics*, vol. 9, no 9, p. 657-662.
- Stewart, J. B., C. Freyer, J. L. Elson, A. Wredenberg, Z. Cansu, A. Trifunovic et N.-G. Larsson. 2008b. «Strong purifying selection in transmission of mammalian mitochondrial DNA». *PLoS Biology*, vol. 6, no 1, p. e10.
- Strayer, D. L. 2006. «Challenges for freshwater invertebrate conservation». *Journal of the North American Benthological Society*, vol. 25, no 2, p. 271-287.
- Subramanian, A., M. Kaufmann et B. Morgenstern. 2008. «Dialign-tx: Greedy and progressive approaches for segment-based multiple sequence alignment». *Algorithms for Molecular Biology*, vol. 3, no 1, p. 6.
- Sutherland, B., D. Stewart, E. R. Kenchington et E. Zouros. 1998. «The fate of paternal mitochondrial DNA in developing female mussels, *Mytilus*

- edulis*: Implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA». *Genetics*, vol. 148, no 1, p. 341-347.
- Sutovsky, P., R. D. Moreno, J. Ramalho-Santos, T. Dominko, C. Simerly et G. Schatten. 1999. «Development: Ubiquitin tag for sperm mitochondria». *Nature*, vol. 402, no 6760, p. 371-372.
- Sutovsky, P., R. D. Moreno, J. Ramalho-Santos, T. Dominko, C. Simerly et G. Schatten. 2000. «Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos». *Biology of Reproduction*, vol. 63, no 2, p. 582-590.
- Swanson, W. J. et V. D. Vacquier. 2002. «Reproductive protein evolution». *Annual Review of Ecology and Systematics*, vol. 33, no 1, p. 161-179.
- Swofford, D. (2002). Paup*. Phylogenetic analysis using parsimony (*and other methods), Sunderland: Sinauer Associates. [<http://paup.csit.fsu.edu>].
- Taanman, J.W. 1999. «The mitochondrial genome: structure, transcription, translation and replication». *Biochimica et Biophysica Acta - Bioenergetics*, vol. 1410, no 2, p. 103-123.
- Tamura, K., Dudley, J., Nei, M. et Kumar, S. 2007. «MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0». *Molecular Biology and Evolution*, vol. 24, p. 1596-1599.
- Taylor, R. W., M. J. Barron, G. M. Borthwick, A. Gospel, P. F. Chinnery, D. C. Samuels, G. A. Taylor, S. M. Plusa, S. J. Needham, L. C. Greaves, T. B. L. Kirkwood et D. M. Turnbull. 2003. «Mitochondrial DNA mutations in human colonic crypt stem cells». *Journal of Clinical Investigation*, vol. 112, no 9, p. 1351-1360.
- Taylor, R. W. et D. M. Turnbull. 2005. «Mitochondrial DNA mutations in human disease». *Nature Reviews Genetics*, vol. 6, no 5, p. 389-402.
- Templeton, A. R., K. A. Crandall et C. F. Sing. 1992. «A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation». *Genetics*, vol. 132, no 2, p. 619-633.

- Theologidis, I., S. Fodelianakis, M. B. Gaspar et E. Zouros. 2008. «Doubly uniparental inheritance (dui) of mitochondrial DNA in *Donax trunculus* (Bivalvia:Donacidae) and the problem of its sporadic detection in Bivalvia». *Evolution*, vol. 62, no 4, p. 959-970.
- Thompson, J. D., D. G. Higgins et T. J. Gibson. 1994. «Clustal w: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice». *Nucleic Acids Research*, vol. 22, no 22, p. 4673-4680.
- Thompson, W. E., J. Ramalho-Santos et P. Sutovsky. 2003. «Ubiquitination of prohibitin in mammalian sperm mitochondria: Possible roles in the regulation of mitochondrial inheritance and sperm quality control». *Biology of Reproduction*, vol. 69, no 1, p. 254-260.
- Thyagarajan, B., R. A. Padua et C. Campbell. 1996. «Mammalian mitochondria possess homologous DNA recombination activity». *Journal of Biological Chemistry*, vol. 271, no 44, p. 27536-27543.
- Timmis, J. N., M. A. Ayliffe, C. Y. Huang et W. Martin. 2004. «Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes». *Nature Reviews Genetics*, vol. 5, no 2, p. 123-135.
- Torrado, M., M. Paz, L. I. Korochkin et A. T. Mikhailov. 2003. «Mussel map, a major gonad-duct esterase-like protein, is released into sea water as a dual constituent of the seminal fluid and the spermatozoon». *Journal of Experimental Biology*, vol. 206, no 2, p. 313-326.
- Trifunovic, A., A. Wredenberg, M. Falkenberg, J. N. Spelbrink, A. T. Rovio, C. E. Bruder, M. Bohlooly-Y, S. Gidlof, A. Oldfors, R. Wibom, J. Tornell, H. T. Jacobs et N. G. Larsson. 2004. «Premature ageing in mice expressing defective mitochondrial DNA polymerase». *Nature*, vol. 429, no 6990, p. 417-423.
- Turgeon, D. D., J. F. Quinn, A. E. Bogan, E. V. Coan, F. G. Hochberg, W. G. Lyons, P. M. Mikkelsen, R. J. Neves, C. F. E. Roper, G. Rosenberg, B. Roth, A. Scheltema, F. G. Thompson, M. Vecchione et J. D. Williams. 1998. *Common and scientific names of aquatic invertebrates from the United States and Canada: Mollusks*. Bethesda, Maryland: American Fisheries Society.

- Turgeon, D.D., A.E. Bogan, E.V. Coan, W.K. Emerson, W.G. Lyons, W.L. Pratt, C.F.E. Roper, A. Scheltema, F.G. Thompson, et J.D. Williams., 1988. *Common and scientific names of aquatic invertebrates from the United States and Canada: Mollusks*. Bethesda, Maryland: American Fisheries Society.
- Turnbull, D. M. et R. N. Lightowlers. 2001. «Might mammalian mitochondria merge?». *Nature Medicine*, vol. 7, no 8, p. 895-896.
- Vallès, Y. et J. L. Boore. 2006. «Lophotrochozoan mitochondrial genomes». *Integrative and Comparative Biology* vol. 46, no 4, p. 544-557.
- Veillette, J. J. 1994. «Evolution and paleohydrology of glacial lakes Barlow and Ojibway». *Quaternary Science Reviews*, vol. 13, no 9-10, p. 945-971.
- Venetis, C., I. Theologidis, E. Zouros et G. C. Rodakis. 2006. «No evidence for presence of maternal mitochondrial DNA in the sperm of *Mytilus galloprovincialis* males». *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 273, no 1600, p. 2483-2489.
- Venetis, C., L. Theologidis, E. Zouros et G. C. Rodakis. 2007. «A mitochondrial genome with a reversed transmission route in the mediterranean mussel *Mytilus galloprovincialis*». *Gene*, vol. 406, no 1-2, p. 79-90.
- Vissing, J. et M. Schwartz. 2002. «Paternal inheritance of mitochondrial DNA. Reply». *New England Journal of Medicine*, vol. 347, no 25, p. 2081-2082.
- Wai, T., D. Teoli et E. A. Shoubridge. 2008. «The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes». *Nature Genetics*, vol. 40, no 12, p. 1484-1488.
- Walker, J. M., J. P. Curole, D. E. Wade, E. G. Chapman, A. E. Bogan, G. T. Watters et W. R. Hoeh. 2006. «Taxonomic distribution and phylogenetic utility of gender-associated mitochondrial genomes in the Unionoida (Bivalvia)». *Malacologia*, vol. 48, no 1-2, p. 265-282.
- Walker, K. F., M. Byrne, C. W. Hickey et D. S. Roper. 2001. «Freshwater mussels (Hyriidae) of Australasia». In *Ecology and evolution of the*

- freshwater mussels Unionida*, G. Bauer et W. Wächtler (éd.), p. 5-31. Berlin, Springer.
- Wallace, D. C. 1992. «Diseases of the mitochondrial DNA». *Annual Review of Biochemistry*, vol. 61, no 1, p. 1175-1212.
- Wallace, D. C. 1999. «Mitochondrial diseases in man and mouse». *Science*, vol. 283, no 5407, p. 1482-1488.
- Wallace, D. C. 2005. «A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: A dawn for evolutionary medicine». *Annual Review of Genetics*, vol. 39, p. 359-407.
- Wang, Y. et X. Guo. 2004. «Chromosomal rearrangement in pectinidae revealed by rRNA loci and implications for bivalve evolution». *Biological Bulletin*, vol. 207, no 3, p. 247-256.
- Wang, Y., Z. Xu et X. Guo. 2005. «Chromosomal mapping of 5s ribosomal RNA genes in the eastern oyster, *Crassostrea virginica* gmelin by fluorescence *in situ* hybridization». *J Shellfish Res*, vol. 24, p. 959 - 964.
- Whelan, N.V., A.J. Geneva et D.L. Graf. 2011. «Molecular phylogenetic analysis of tropical freshwater mussels (Mollusca: Bivalvia: Unionoida) resolves the position of *Coelatura* and supports a monophyletic Unionidae». *Molecular Phylogenetics and Evolution*, vol. 61, p. 504-514.
- Whelan, S. et N. Goldman. 2001. «A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach». *Molecular Biology and Evolution*, vol. 18, no 5, p. 691-699.
- White, D. J., Wolff, J. N. , Pierson, M. et N.J. Gemmell. 2008. «Revealing the hidden complexities of mtDNA inheritance». *Molecular Ecology*, vol. 17, no 23, p.4925-4942.
- Wiens, J. J. et T. A. Penkrot. 2002. «Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*)». *Systematic Biology*, vol. 51, no 1, p. 69-91.

- Williams, J. D., M. L. Warren, K. S. Cummings, J. L. Harris et R. J. Neves. 1993. «Conservation status of freshwater mussels of the United States and Canada». *Fisheries*, vol. 18, no 9, p. 6-22.
- Williams, R. S. 2002. «Another surprise from the mitochondrial genome». *New England Journal of Medicine*, vol. 347, no 8, p. 609-612.
- Wilson, D. M. et V. A. Bohr. 2007. «The mechanics of base excision repair, and its relationship to aging and disease». *DNA Repair*, vol. 6, no 4, p. 544-559.
- Wolstenholme, D. R. 1992. «Animal mitochondrial DNA: structure and evolution». *International Review of Cytology*, vol. 141, p. 173-216.
- Wood, A. R., G. Turner, D. O. F. Skibinski et A. R. Beaumont. 2003. «Disruption of doubly uniparental inheritance of mitochondrial DNA in hybrid mussels (*Mytilus edulis* x *M. galloprovincialis*)». *Heredity*, vol. 91, no 4, p. 354-360.
- Wyman, S. K., R. K. Jansen et J. L. Boore. 2004. «Automatic annotation of organellar genomes with dogma». *Bioinformatics*, vol. 20, no 17, p. 3252-3255.
- Xu, J. P. 2005. «The inheritance of organelle genes and genomes: Patterns and mechanisms». *Genome*, vol. 48, no 6, p. 951-958.
- Xu, X., X. Wu et Z. Yu. 2010. «The mitogenome of *Paphia euglypta* (Bivalvia: Veneridae) and comparative mitogenomic analyses of three venerids». *Genome*, vol. 53, no 12, p. 1041-1052.
- Yaffe, M. P. 1999. «Dynamic mitochondria». *Nature Cell Biology*, vol. 1, no 6, p. E149-E150.
- Yang, Z. 1993. «Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods». *Molecular Biology and Evolution*, vol. 39, no 3, p. 306-314.
- Yang, Z. 1998. «Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution». *Molecular Biology and Evolution*, vol. 15, no 5, p. 568-573.

- Yang, Z. 2002. «Inference of selection from multiple species alignments». *Current Opinion in Genetics and Development*, vol. 12, no 6, p. 688.
- Yang, Z. 2007. «Paml 4: Phylogenetic analysis by maximum likelihood». *Molecular Biology and Evolution*, vol. 24, no 8, p. 1586-1591.
- Yang, Z. et R. Nielsen. 2008. «Mutation-selection models of codon substitution and their use to estimate selective strengths on codon usage». *Molecular Biology and Evolution*, vol. 25, no 3, p. 568-579.
- Yang, Z., W. S. W. Wong et R. Nielsen. 2005. «Bayes empirical bayes inference of amino acid sites under positive selection». *Molecular Biology and Evolution*, vol. 22, no 4, p. 1107-1118.
- Yang, Z. H., R. Nielsen, N. Goldman et A. M. K. Pedersen. 2000. «Codon-substitution models for heterogeneous selection pressure at amino acid sites». *Genetics*, vol. 155, no 1, p. 431-449.
- Yasukawa, T., M.-Y. Yang, H. T. Jacobs et I. J. Holt. 2005. «A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA». *Molecular cell*, vol. 18, no 6, p. 651-662.
- Yu, Z., Z. Wei, X. Kong et W. Shi. 2008. «Complete mitochondrial DNA sequence of oyster *Crassostrea hongkongensis* — a case of "Tandem duplication-random loss" for genome rearrangement in *Crassostrea*?». *BMC Genomics*, vol. 9, no 477, doi : 10.1186/1471-2164-9-477.
- Yusa, Y. 2007. «Causes of variation in sex ratio and modes of sex determination in the Mollusca — an overview». *American Malacological Bulletin*, vol. 23, no 1-2, p. 89-98.
- Zbawicka, M., A. Burzynski et R. Wenne. 2007. «Complete sequences of mitochondrial genomes from the baltic mussel *Mytilus trossulus*». *Gene*, vol. 406, no 1-2, p. 191-198.
- Zeh, J. A. 2004. «Sexy sons: A dead end for cytoplasmic genes». *Proceedings of the Royal Society of London Series B-Biological Sciences*, vol. 271, p. S306-S309.
- Zeh, J. A. et D. W. Zeh. 2005. «Maternal inheritance, sexual conflict and the maladapted male». *Trends in Genetics*, vol. 21, no 5, p. 281-286.

- Zeviani, M., P. Amati, G. Comi, G. Fratta, C. Mariotti et V. Tiranti. 1995. «Searching for genes affecting the structural integrity of the mitochondrial genome». *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1271, no 1, p. 153-158.
- Zhang, J. 2003. «Evolution by gene duplication: An update». *Trends in Ecology and Evolution*, vol. 18, no 6, p. 292-298.
- Zhang, J. Z., R. Nielsen et Z. H. Yang. 2005. «Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level». *Molecular Biology and Evolution*, vol. 22, no 12, p. 2472-2479.
- Zink, R. M. et G. F. Barrowclough. 2008. «Mitochondrial DNA under siege in avian phylogeography». *Molecular Ecology*, vol. 17, no 9, p. 2107-2121.
- Zorzano, A., M. Liesa, D. Sebastián, J. Segalés et M. Palacín. 2010. «Mitochondrial fusion proteins: Dual regulators of morphology and metabolism». *Seminars in Cell and Developmental Biology*, vol. 21, no 6, p. 566-574.
- Zouros, E. 2000. «The exceptional mitochondrial DNA system of the mussel family Mytilidae». *Genes and Genetic Systems*, vol. 75, no 6, p. 313-318.
- Zouros, E., A. O. Ball, C. Saavedra et K. R. Freeman. 1994a. «Mitochondrial-DNA inheritance - reply». *Nature*, vol. 368, no 6474, p. 818-818.
- Zouros, E., A. O. Ball, C. Saavedra et K. R. Freeman. 1994b. «An unusual type of mitochondrial-DNA inheritance in the blue mussel *Mytilus*». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no 16, p. 7463-7467.
- Zouros, E., K. R. Freeman, A. O. Ball et G. H. Pogson. 1992. «Direct evidence for extensive paternal mitochondrial-DNA inheritance in the marine mussel *Mytilus*». *Nature*, vol. 359, no 6394, p. 412-414.
- Zwickl, D. (2006). Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion, *PhD thesis*. The University of Texas at Austin.